Scientific and Technical Information Center

SEARCH REQUEST FORM

Date: _1/10/00_ Requester's Full Name: _____ Examiner #: _Li Lee_

Art Unit: _1645_ Phone (30_8_) _6891_ Serial Number: _09/041,236_

Results Format Preferred (circle): PAPER DISK E-MAIL

****

To ensure an efficient and quality search, please attach a copy of the cover sheet, claims, and abstract or fill out the following:

Title of Invention: _____

Inventors (please provide full names): _____

_____

Earliest Priority Date: _____

Search Topic:
Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known.

For Sequence Searches Only Please include all pertinent information (parent, grandchild, divisional, or issued patent numbers) along with the appropriate serial number.

please search =

R.   (NA)

Back translate SEQ ID NO: 2 into NA and run against NA datebase
(AA)

Point of Contact:
Susan Hanley
Technical Info. Specialist
CM1 12C14 Tel: 305-4053

FOR OFFICIAL USE ONLY

****

STAFF USE ONLY	Type of Search	Vendors and Cost	
Searcher: _____	_____ NA Sequence (#)	_____ STN	_____ Dialog
Searcher Phone #: _____	_____ AA Sequence (#)	_____ Questel/Orbit	_____ Dr.Link
Searcher Location: _____	_____ Structure (#)	_____ Lexis/Nexis	_____ Westlaw
Date Searcher Picked Up: _____	_____ Bibliographic	_____ WWW/Internet	
Date Completed: _____	_____ Litigation	_____ In-house sequence systems (list)	
Searcher Prep & Review Time: _____	_____ Fulltext	_____ Other (specify)	
Online Time: _____	_____ Other		

OM of: US-09-041-236-2 to: GenEmbl:*   out_format : pfs
Date: Jan 18, 2000   6:57 PM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-1998 Compugen Ltd.

Command line parameters:
 -MODEL=frame+ p2n.model -DEV=xlp
 -Q=/cgn2_1/USPTO_spool/US09041236/runat_15012000_164051_503/app_query.fasta.1
 -DB=GenEmbl -QFMT=fastap -SUFFIX-LEE.rge -GAPOP=12.000
 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
 -OGAPOP=4.500 -OGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.000
 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
 -DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
 -MAXLEN=1000000 -USER=US09041236 -NCPU=6 -ICPU=3 -NO_XLPXY -WAIT
 -THREADS=1

Search information block:
Query: US-09-041-236-2
Query length: 634
Database: GenEmbl:*
Database sequences: 780561
Database length: 213795050
Search time (sec): 539.620000

```
seq_name: gb_pr3:AF030698

                                                  +   535.50   745.42   1.6e-33   3988   ! Z93948 M.musculus mRNA for
gb_ro:MM293948                                    +   525.50   733.99   7.0e-33   2993   ! AF073289 Danio rerio transm
gb_nv:AF073289                                    +   500.50   695.06   1.0e-30   4002   ! AB021291 Mus musculus mRNA
gb_o:AB021291                                     +   489.50   679.23   7.8e-30   4008   ! AB002563 Rattus norvegicus
gb_ro:AB002563
```

seq_documentation_block:
```
LOCUS       AF030698          2636 bp    mRNA             PRI      03-SEP-1998
DEFINITION  Homo sapiens semaphorin L (SEMAL) mRNA, complete cds.
ACCESSION   AF030698
NID         g3523114
VERSION     AF030698.1  GI:3523114
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2636)
  AUTHORS   Lange,C., Liehr,T., Goen,M., Gebhart,E., Fleckenstein,B. and
            Ensser,A.
  TITLE     New eukaryotic semaphorins with close homology to semaphorins of
            DNA viruses
  JOURNAL   Genomics 51 (3), 340-350 (1998)
  MEDLINE   98389619
REFERENCE   2  (bases 1 to 2636)
  AUTHORS   Ensser,A., Liehr,T., Lange,C., Goen,M., Gebhart,E. and
            Fleckenstein,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-OCT-1997) Institut fuer Klinische und Molekulare
            Virologie, Friedrich Alexander Universitaet, Schlossgarten 4,
            Erlangen, Bavaria 91054, Germany
FEATURES             Location/Qualifiers
     source          1..2636
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /map="15p22.3-p23"
                     /chromosome="15"
     gene            1..2636
                     /gene="SEMAL"
     CDS             13..2013
                     /gene="SEMAL"
                     /note="H-Sema-L"
                     /codon_start=1
                     /product="semaphorin L"
                     /protein_id="AAC34261.1"
                     /db_xref="PID:g3523115"
                     /db_xref="GI:3523115"
                     /translation="MTPPPGRAAPSAPRARVPGPPARLGLPLRLRLLLLWAAASA
                     QGHLRSGPRIFAVWKGHVGQDRVDFGQTEPHTVLFHEPGSSSVWVGGRGKVYLFDFPE
                     GKNASVRRTVNIGSTKGSCLDKRDCENYITLLERRSEGLLACGTNARHPSCWNLVNGTV
                     VPLGEMRGYAPFSPDENSLVLFEGDEVYSTIRKQEYNGKIPRFRIRGESELYTSDTV
                     MQNPQFIKATIVHDQAYDDKIYFFREDNPDKNPEAPLNVSRVAQLCRGDQGGSSL
                     SVSKWNTFLKAMLVCSDAATNKNFNRLQDVFLLPDSGGWRDTRVYGVFSNPWNYSAV
                     CVFSLGDIDKVFRTSSLKGYHSSLPNPRPGKCLPDQQPIPTETQVADRHPEVAQRVE
                     PMGPLKTPLFHSKYHYQKVAVHRMQASHGETFHVLYLTTDRGTIHKVVEPGEQHSRD
                     FNIMEIQPFRRAAAIQTMSLDAERRKLYYSSQWEVSQVPLDLCEVYGGGCHGCLMSRD
                     PYCGWDQGRCISIYSSERSVLQSINPAEPHKECPNPKPDKAPLQKVSLAPNSRYLSC
                     PMESRHATYSWRHKENVEQSCEPGHQSPNCILFIENLTAQQYGHYFCEAQEGSYPREA
                     QHWQLLPEDGIMAEHLLGHACALAASLWLGVLPTLTLGLLVH"
BASE COUNT      536 a    835 c    764 g    501 t
ORIGIN
```

alignment_scores:
```
       Quality: 3443.00              Length:      634
         Ratio: 5.431                  Gaps:        0
Percent Similarity: 100.000   Percent Identity: 100.000
```

alignment_block:
US-09-041-236-2 x AF030698

sgn seq 1/1 to: AF030698   from: 1 to: 2636

score_list:
```
Sequence                Strd    Orig      ZScore    EScore  Len   ! Documentation
gb_pr3:AF030698           +   3443.00   4930.07   1.3e-266  2636  ! AF030698 Homo sapiens semaphor
gb_pr4:AF069493           +   3443.00   4929.97   1.3e-266  2661  ! AF069493 Homo sapiens GPI-anch
gb_pr4:AF071542           +   3443.00   4930.23   1.3e-266  2594  ! AF071542 Homo sapiens semaphor
gb_ro:AB017532            +   3072.00   4396.76   6.7e-237  2608  ! AB017532 Mus musculus mRNA fcr
gb_ro:AF176670            +   3072.00   4397.04   6.5e-237  2538  ! AF176670 Mus musculus GPI-anch
gb_pr3:AF030697           +   1907.50   2709.86   6.1e-143  8861  ! AF030697 Homo sapiens semaphor
gb_pr3:AF030699           +   1748.00   2501.16   2.6e-131  1192  ! AF030699 Mus musculus semaphor
gb_v1:AF005370            +   1342.00   1869.13   4.1e-96  130608  ! AF005370 Alcelaphine herpesv
gb_v1:AHU18243            +   1342.00   1893.78   1.7e-97  11874  ! U18243 Alcelaphine herpesviru
gb_pat:AR040598           +    649.50    913.72   6.8e-43   2601  ! AR040598 Sequence 53 from pate
gb_pat:I47054             +    649.50    913.72   6.8e-43   2601  ! I47054 Sequence 53 from patent
gb_pr2:HUMHSEM            +    649.50    914.00   6.6e-43   2530  ! L26081 Homo sapiens semaphorin
gb_sts:G31703             +    649.50    914.00   6.6e-43   2530  ! G31703 sWSS1973 Eric D. Green
gb_ro:MUSC1               +    635.00    893.89   8.7e-42   2356  ! D85028 Mouse collapsin-1 mRNA,
gb_ro:MRRNASEMA           +    634.50    891.13   1.2e-41   2872  ! X95990 R.norvegicus mRNA for sem
gb_ro:MRRNASEMD           +    631.50    885.51   2.5e-41   2319  ! X95286 M.musculus mRNA for s
gb_ov:GGU02528            +    623.00    876.80   7.8e-41   3263  ! U02528 Gallus gallus collapsin
gb_ro:RNSIIICN1           +    621.00    871.58   1.5e-40   2913  ! X85992 M.musculus mRNA for sem
gb_ro:MRRNASEMC           +    617.00    867.68   2.5e-40   2433  ! X85992 M.musculus mRNA for sem
gb_ro:AF083382            +    613.00    862.34   4.0e-40   2337  ! AF083382 Danio rerio semaphori
gb_ro:MUSSEMAIII          +    612.00    862.51   4.0e-40   1998  ! L40484 Mus musculus semaphorin
gb_pr4:HSU28369           +    612.00    858.62   8.0e-40   2919  ! U28369 Homo sapiens semaphorin
gb_ro:HSU38276            +    611.00    855.16   1.0e-39   3552  ! U38276 Human semaphorin III fa
gb_ov:AF086761            +    607.50    851.37   2.0e-39   3148  ! AF086761 Danio rerio semaphori
gb_ro:MRRNASEME           +    597.50    839.46   9.3e-39   2477  ! X85994 M.musculus mRNA for sem
gb_ro:HSU33920            +    597.50    837.78   1.2e-38   2719  ! U33920 Human clone lambda 5 se
gb_ov:AF124485            +    595.50    835.03   1.6e-38   2879  ! AF124485 Danio rerio secreted
gb_pr1:AB000220           +    591.50    823.25   7.4e-38   5177  ! AB000220 Homo sapiens mRNA for
gb_pat:E16310             +    591.00    828.96   3.6e-38   2769  ! E16310 Murine mRNA for semapho
gb_ro:MMU69535            +    591.00    828.96   3.6e-38   2769  ! U69535 Mus musculus semaphorin
gb_ro:AF080091            +    589.50    828.87   3.6e-38   2265  ! AF080091 Mus musculus semaphor
gb_ro:AF080090            +    583.00    819.01   1.5e-37   2383  ! AF080090 Mus musculus semaphor
gb_ov:AF022946            +    575.00    807.71   5.5e-37   2336  ! AF022946 Gallus gallus collaps
gb_pr3:HSU60800           +    570.00    794.59   2.9e-36   4157  ! U60800 Human semaphorin (CD100
gb_ro:AB002329            +    567.50    786.44   8.4e-36   6474  ! AB002329 Human mRNA for KIAA03
gb_ro:AF034744            +    566.00    792.55   3.0e-36   2898  ! AF034744 Mus musculus semaphor
gb_ov:GGU28240            +    564.50    791.07   4.6e-36   2715  ! U28240 Gallus gallus collapsin
gb_ro:MMSEMH              +    555.00    778.99   2.2e-35   2328  ! Z80941 M.musculus mRNA for sem
gb_pat:I11900             +    547.00    772.14   5.2e-35   1481  ! I11900 Sequence 1 from Patent.
gb_ov:AF022947            +    538.50    754.07   5.3e-34   2615  ! AF022947 Gallus gallus collaps
gb_ro:MMZ93947            +    535.50    744.25   1.9e-33   4466  ! Z93947 M.musculus mRNA for sem
```

```
 301 AlaValCysValTyrSerLeuGlyAspIleAspLysValPheArgThrSe  317
1009 GCCGTCTGTGTATTCCCTCGGTAACATTGACAAGGTCTTCCGTACCTC    1058

 317 rSerLeuLysGlyTyrHisSerSerLeuProAsnProArgProGlyLysC  334
1059 CTCACTCAAGGGCTACCACTCAAGCCTTCCCAACCCGCGGCCTGGCAAGT  1108

 334 ysLeuProAspGlnProIleProThrGluThrPheGlnValAlaAsp     350
1109 GCCTCCCAGACCAGCAGCGATACCCACAGAGACCTTCCAGGTGGCTGAC   1158

 351 ArgHisProGluValAlaGlnValGluProMetGlyProLeuLysTh    367
1159 CGTCACCCAGAGGTGGCCCAGGTGGAGCCCATGGGCCCTCTGAAGAC     1208

 367 rProLeuPheHisSerLysTyrHisTyrGlnLysLysValAlaValHisArgM 384
1209 GCCATTGTTCCACTCTAAATACCACTACCAGAAAGTGGCCGTTCACCGCA  1258

 384 etGlnAlaSerHisGlyGluThrPheHisValLeuTyrLeuThrThrAsp  400
1259 TGCAAGGCCAGCCACGGGGAGACCTTTCATGTGCTTTACCTAACTACAGAC 1308

 401 ArgGlyThrIleHisLysValValGlnProGlyGluGlnHisGlnHisSerPh 417
1309 AGGGGCACTATCCACAAGGTGGTGCAACCGGGGGAGCAGGAGCACAGCTT  1358

 417 eAlaPheAsnIleMetGluIleGlnProPheArgArgArgAlaAlaAlaIleG 434
1359 CGCCTTCAACATCATGGAGATCCAGCCCTTCCGCCGGCGGCTGCCATCC   1408

 434 lnThrMetSerLeuAspAlaGluArgLysLeuArgArgValSerSerGln  450
1409 AGACCATGTCGCTGGATGCTGAGCGCGGAGGAAGCTGTATGTGAGCTCCCAG 1458

 451 TrpGluValSerGlnValProLeuAspLeuCysGluValTyrGlyValGlyI 467
1459 TGGGAGGTGAGCCAGGTGCCCCTGGACCTGTGTGAGGTCTATGGCGGGGG  1508

 467 yCysHisGlyCysLeuMetSerArgAspProTyrCysGlyTrpAspGlnG  484
1509 CTGCCACGGTTGCCTCATGTCCCGAGACCCCTACTGCGGCTGGGACCAGG  1558

 484 lyArgCysIleSerIleTyrSerSerGluArgSerValLeuGlnSerIle  500
1559 GCCGCTGCATCTCCATCTACAGCTCCGAACGGTCAGTGCTGCAATCCATT  1608

 501 AsnProAlaGluProHisLysGluCysProAsnProLysProAspLysVAl 517
1609 AATCCAGCCGAGCCACACAAGGAGTGTCCCAACCCCAAACCAGACAAGGC  1658

 517 aProLeuGlnLysValSerLeuAlaProAsnSerArgArgTyrTyrLeuSerC 534
1659 CCCACTGCAGAAGGTTTCCCTGGCCCCAAACTCTCGCTACTACCTGAGCT  1708

 534 ysHisProMetGluSerArgHisAlaThrTyrSerTrpArgHisLysGluAsn 550
1709 GCCCCATGGAATCCCGGCACGCCACCTACTCATGGCGCCACAAGGAGAAC  1758

 551 ValGluGlnSerCysGluGluCysProGlyHisGlnSerProAsnCysIleLeuPh 567
1759 GTGGAGCAGAGCTGCGAAGAGTGCCCTGGTCACCAGAGCCCCAACTGCATCCTGTT 1808

 567 eIleGluAsnLeuThrAlaGlnTyrGlyHisTyrPheCysGluAlaG   584
1809 CATCGAGAACCTTACGGCGCAGTACGGCCACTACTTCTGCGAGGCCC     1858

 584 lnGluGlySerTyrPheArgGluAlaGlnHisTrpLeuLeuProGlu   600
1859 AGGAGGGCTCCTACTTCCGGGAGGCTCAGCACTGGCTGCTGCCGGAG    1908

 601 AspGlyIleMetAlaGluHisHisLeuGlyHisHisAlaCysAlaLeuAlaAl 617
```

```
   1 LeuLeuLeuLeuTrpAlaAlaAlaAlaSerArgAlaGlnAlaGlyHisLeuAr  17
 109 CTGCTGCTGCTGTGGGCGGCCGCCGCTCCGGGCGCCCAGGGCCACTTAAG    158

  17 gSerGlyProArgIlePheAlaValTrpLysGlyHisGlnValGlyAspA    34
 159 GAGCGGACCCCGCATCTTCGCCGTCTGGAAAGGCCATGTAGGGCAGGACC   208

  34 rgValAspPheGlyGlnThrGluProHisThrValLeuPheHisGluPro    50
 209 GGGTGGACTTTGGCCAGACTGAGCCGCACACGGTGCTTTTCCACGAGCCA   258

  51 GlySerSerValTrpValGlyGlyLysValTyrLeuPheAs             67
 259 GGCAGCTCCGTCTGGGTGGGTGGGAAGGTCTACCTCTTTGA            308

  67 pPheProGluGlyLysAsnAlaSerValArgThrValAsnIleGlySerT    84
 309 CTTCCCAGAGGGCAAGAACGCATCTGTGCGCACGGTGAATATCGGCTCCA   358

  84 hrLysGlySerCysLeuAspLysAspCysSerGluAsnTyrIleThrLeu  100
 359 CAAAGGGGTCCTGTCTGGATAAAGACTGCGAGAACTACATCACTCTC     408

 101 LeuGluArgArgSerGluGlyLeuLeuLeuAlaCysGlyThrAsnAlaArgHi 117
 409 CTGGAGAGGCGGAGTGAGGGGCTGCTGCTGGCCTGCGGCACCAACGCCCGGCA 458

 117 sProSerCysTrpAsnLeuValAsnGlyThrValProLeuGlyGluM     134
 459 CCCCAGCTGCTGGAACCTGGTGAATGGCACTGTGCCACTTGGCGAGA    508

 134 etArgGlyTyrTyrAlaProPheSerProAspGluAsnSerLeuValLeuPhe 150
 509 TGAGAGGCTACTACGCCCCCTTCAGCCCCGACGAGAACTCCCTGGTTCTGTTT 558

 151 GluGlyAspGluValTyrSerThrIleArgLysGlnLeuTyrAsnGlyLy  167
 559 GAAGGGGACGAGGTGTATTCCACCATCCGGAAGCAGCAGAATACAATGGGAA 608

 167 sIleProArgPheArgArgIleArgGlyArgGluSerGluLeuTyrThrSerA 184
 609 GATCCCTCGGTTCCGCCGCATCCGGGGCCGAGAGTGAGCTGTACACCAGTG  658

 184 spThrValMetGlnAsnProGlnPheIleLysValAlaThrIleValHisGln 200
 659 ATACTGTCATGCAGAACCCACAGTTCATCAAAGCCACCATCGTGCACCAA   708

 201 AspGlyGlnLeuAlaTyrAspAspLysIleTyrTyrPhePheArgGluAspAsnPr 217
 709 GACCAGGCTTACGATGACAAGATCTACTACTTCTTCCGAGAGGACAATCC  758

 217 oAspLysAsnProGluAlaProLeuAsnAlaSerArgValAlaGlnLeuC   234
 759 TGACAAGAATCCTGAGGCTCCCCTCAATGTGCCCGTGGCCCAGTTGT     808

 234 ysArgGlyAspGlnGlyGlyGlySerLeuSerLysLysSerTrpAsn      250
 809 GCAGGGGGGACCAGGGTGGGGGAAGTTCACTCTGTCTCCAAGTGGAAC    858

 251 ThrPheLeuLysMetAlaMetLeuValCysSerAspSerAlaThrAsnLysAs 267
 859 ACTTTTCTGAAAGCCATGGCCATGCTGGTATGCAGTGATTCTGCCACCAACAAGAA 908

 267 nPheAsnArgLeuGlnAspValPheLeuLeuProAspProSerGlyGlnT    284
 909 CTTCAACAGGCTGCAAGACGTCTTCCTGCTCCCTGACCCCAGCGGCCAGT   958

 284 rpArgAspThrArgValThrArgValGlyValPheSerAsnProTrpAsnTyrSer 300
 959 GGAGGGACACCAGGGTCTATGGTGTTTTCTCCAACCCCTGAACTACTCA   1008
```

```
 1909 GACGGCATCATGGCCCAGCACCTGGGTCATGGCTGCCCTGGCTGC 1958
      ||||||||||||||||||||||||||||||||||||||||||||||
  617 aSerLeuTrpLeuGlyValLeuProThrLeuThrLeuGlyLeuLeuValH 634
      ||||||||||||||||||||||||||||||||||||||||||||||
 1959 CTCCCTCTGCTGGGGGTGCTGCCCACACTCACTCTTGGCTTGCTGGTCC 2008

  634 1s 634
      ||
 2009 AC 2010

seq_name: gb_pr4:AF069493

seq_documentation_block:
LOCUS       AF069493    2661 bp    mRNA         PRI    28-MAY-1999
DEFINITION  Homo sapiens GPI-anchored membrane protein CDw108 precursor, mRNA,
            complete cds.
ACCESSION   AF069493
NID         g3947939
VERSION     AF069493.1    GI:3947939
KEYWORDS    .
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2661)
  AUTHORS   Yamada,A., Kubo,K., Takeshita,T., Harashima,N., Kawano,K., Mine,T.,
            Sagawa,K., Sugamura,K. and Itoh.K.
  TITLE     Molecular cloning of a glycosylphosphatidylinositol-anchored
            molecule CDw108
  JOURNAL   J. Immunol. 162 (7), 4094-4100 (1999)
  MEDLINE   99178431
REFERENCE   2  (bases 1 to 2661)
  AUTHORS   Yamada.A., Kubo,K., Takeshita,T., Harashima,N., Kawano,K.,
            Sagawa,K., Sugamura,K. and Itoh,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1998) Immunology, Kurume University School of
            Medicine, Asahi-machi 67, Kurume, Fukuoka 830, Japan
FEATURES             Location/Qualifiers
     source          1..2661
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="15"
                     /map="between WI-6247 and CHLC.GATA85D02"
                     /cell_line="HPB-ALL"
                     /cell_type="T lymphocyte"
     CDS             18..2018
                     /note="putative adhesion molecule; lymphocyte activation
                     antigen"
                     /codon_start=1
                     /product="GPI-anchored membrane protein CDw108 precursor"
                     /protein_id="AAC82642.1"
                     /db_xref="PID:g3947940"
                     /db_xref="GI:3947940"
                     /translation="MTPPPGRAAPSAPRARVPGPPARLGLPLRLRLLLLWAAAASA
                     QGHLRSGPRIFAVWKGHVGQDRVDFGQTEPHTVLFHEPGSSSVWVGGRGKVYLFDFPE
                     GKNASVRTVNIGSTKGSCLDKRDCENYITLLERRSEGLLACGTNARHPSCWNLVNGTV
                     VPLGEMRGYAPFSPDENSLVLFEGDEVVSTIRQEYNGKIPRFRRIRGESELYTSDTV
                     MQNPOFIKATIVHQDQAYDDKIYYFFREDNPDKNPEAPLNVSRVAQLCRGDQGESSL
                     SVSKWNTFLKAMLVCSDAATNKNFNRLQDVFLLPDPSGQWRDTRVYGVFSNPWNYSAV
                     CVYSLGDIDKVFRTSSLKGYHSSLPNPRPGKCLPDQQPIPTEFQVADRHPEVAQRVE
                     PMGPLKTPLFHSKYHYQKVAVHRMQASHGETFHVLYLITDRGTIHKVVEPGEQEHSFA
                     FNIMEIQPFRRAAAIQTMSLDAERRKLIVSSQWEVSQVPLDLCEVYGGGCHGCLMSRD
                     PYCGWDQGRCISIYSSERSVLQSINPAEPHKECPNPKFPDKAPLQKVSLAPNSRYYLSC
                     PMESRHATSWRHKENVEQSCEPGHQSPNCILFIENLTAQQYGHYFCEAQEGSYFREA
                     QHWQLLPEDGIMAEHLLGHACALAASLWLGVLPTLTGLLVH"
     sig_peptide     18..155
     misc_feature    156..212
                     /note="Region: N-terminal peptide"
     mat_peptide     156..1961
                     /product="GPI-anchored membrane protein CDw108"
     misc_feature    330..341
                     /note="glycosylation site"
     misc_feature
                     /evidence=not_experimental
                     486. .497
                     /note="glycosylation site"
                     /evidence=not_experimental
     misc_feature    789. .800
                     /note="glycosylation site"
                     /evidence=not_experimental
     misc_feature    816. .824
                     /evidence=not_experimental
                     /note="Region: RGD cell attachment sequence"
     misc_feature    1005. .1016
                     /note="glycosylation site"
                     /evidence=not_experimental
     misc_feature    1821. .1832
                     /note="glycosylation site"
                     /evidence=not_experimental
     misc_feature    1959. .2015
                     /note="Region: GPI-anchor motif"
BASE COUNT      522 a    850 c    769 g    520 t
ORIGIN

alignment_scores:
                     Quality:  3443.00          Length:      634
                     Ratio:    5.431            Gaps:        0
Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-041-236-2 x AF069493

Align seq 1/1  to: AF069493  from: 1  to: 2661

    1 LeuLeuLeuLeuProAlaAlaAlaAlaAlaSerAlaGlnGlyHisLeuAr  17
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  114 CTGCTGCTGCTCCCTGGGCGGCCGCCTCCGCCCCAGGGCCACCTAAG   163

   17 gSerGlyProArgIlePheAlaValTrpLysGlyHisValGlyGlnAspA  34
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  164 GAGCGGACCCCGCATCTTCGCCGTCTGAAAGGCCATGTAGGGCAGGACC   213

   34 rgValAspPheGlyGlnThrGluProHisThrValLeuPheHisGluPro  50
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  214 GGGTGGACTTTGGCCAGACTGAGCCGCACACGGTGCTTTTCCACGAGCCA   263

   51 GlySerSerSerValTrpValGlyGlyArgGlyLysValTyrLeuPheAs  67
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  264 GGCAGCTCCCTCTGTCGTGGGTGGGAGGACGTGGCAAGGTCTACCTCTTGA   313

   67 pPheProGluGlyLysAsnAlaSerValArgThrValAsnIleGlySerT  84
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  314 CTTCCCCGAGGGCAAGAACGCATCTGCGCCACGGTGAATATCGGCTCCA   363

   84 hrLysGlySerCysLeuAspLysArgAspCysGluAsnTyrIleThrLeu  100
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  364 CAAAGGGGTCCTGTCTGGATAAGCGGGACTGCCAGAACTACATCACTCTC   413

  101 LeuGluArgArgSerGluGlyLeuLeuAlaCysGlyThrAsnAlaArgH1  117
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  414 CTGGAGAGGCGGAGTGAGGGGCTGCTGGCCTGTGGCACCAACGCCCCGCA   463

  117 sProSerCysTrpAsnLeuValAsnGlyThrValValProLeuGlyGluM  134
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  464 CCCCAGCTGCTGGAACCTGGTCAATGGCACTGTGGTGCCACTGGGCGAGA   513

  134 etArgGlyTyrAlaProPheSerProAspGluAsnSerLeuValLeuPhe  150
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  514 TGAGAGGCTACGGCCCCCTTCAGCCCCGGACGAGGAGAACTCCCTGGTTCGTTT   563

  151 GluGlyAspGluValTyrSerThrIleArgGlnGluTyrAsnGlyLy  167
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  564 GAAGGGGACGAGGTGTATTCCACCATCCGGAAGCAGGAATACAATGGGAA   613

  167 sIleProArgPheArgArgIleArgGlySerGluLeuTyrThrSerA  184
      ||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
614  GATCCCTCGGTTCCGCCCATCCGGGGCGAGTGAGCTGTACACCAGTG  663
184  spThrValMetGlnAsnProGlnPheIleLysAlaThrIleValHisGln  200
664  ATACTGTCATGCAGAACCACACTTCATCAAAGCCACCATCGTGCACCAA  713
201  AspGlnAlaTyrAspAspLysIleTyrPhePheArgGluAspAsnPr  217
714  GACCAGGCTTACGATGACAAGATCTACTACTTCTTCCGAGAGGACAATCC  763
217  oAspLysAsnProGluAlaProLeuAsnValSerArgValAlaGlnLeuC  234
764  TGACAAGAATCCTGAGGCTCCTCAATGTGTCCCGTGTGGCCCAGTTGT  813
234  ysArgGlyAspGlnGlyGluSerSerLeuSerValSerLysTrpAsn  250
814  GCAGGGGGACCAGGGTGGGGAAAGTTCACTGTCAGTCTCCAAGTGGAAC  863
251  ThrPheLeuLysAlaMetLeuValCysSerAspAlaAlaThrAsnLysAs  267
864  ACTTTTCTGAAAGCCATGCTGGTATGCAGTGATGCTGCCACCAACAAGAA  913
267  nPheAsnArgLeuGlnAspValPheLeuLeuProAspSerGlyGlnT  284
914  CTTCAACAGGCTGCAAGACGTCTTCCTGCTCCCTGACCCCAGCGGCCAGT  963
284  rpArgAspThrArgValTyrGlyValPheSerAsnProTrpAsnTyrSer  300
964  GGAGGGACACCAGGGTCTATGGTGTTTTCTCCAACCCCTGGAACTACTCA  1013
301  AlaValCysValTyrSerLeuGlyLeuAspAspLysPheArgThrSe  317
1014 GCCGTCTGTGTGTATTCCCTCGGTGACAAGGTTCTTCCGTACCTC  1063
317  rSerLeuLysGlyTyrHisSerSerLeuProAsnProArgProGlyLysC  334
1064 CTCACTCAAGGGCTACCACTCAAGCCTTCCCAACCCGCGGCCTGGCAAGT  1113
334  ysLeuProAspGlnProIleProThrGluProThrPheGlnValAlaAsp  350
1114 GCCTCCCAGACCAGCAGCCGCGATCACCCAGAGACCTTCCAGGTGGCTGAC  1163
351  ArgHisProGluAlaAlaGlnArgValGluProMetGlyProLeuLysTh  367
1164 CGTCACCCAGAGGTGGCGCGAGAGGGTGGAGCCCATGGGGCCTCTGAAGAC  1213
367  rProLeuPheHisSerLysTyrHisTyrGlnLysLysValAlaValHisArgM  384
1214 GCCATTGTTCCACTCTAAATACCACTACCAGAAGAAGTGGCGGTCCACCGCA  1263
384  etGlnAlaSerHisSerGlyThrPheHisSerValLeuTyrLeuThrThrAsp  400
1264 TGCAAGGCCAGCCACGCGGGGAGACCTTCATGTGCTTACCTAACTACAGAC  1313
401  ArgGlyThrIleHisLysValValGluValGluGlnGluGlnHisSerPh  417
1314 AGGGGCACTATCCACAAGGTGGTGGAACCGGGGGAGCAGGAGCACAGCTT  1363
417  eAlaPheAsnIleMetGluIleGlnProPheArgArgAlaAlaAlaIleG  434
1364 CGCCTTCAACATCATGGAGATCCAGCCCTTCCGCCGCGCGGCGGCCATCC  1413
434  lnThrSerLeuAspAlaGluArgArgLysTyrValSerSerGln  450
1414 AGACCATGTCGCTGGATGCTGAGCGGAGAAGCTGTATGTGAGCTCCCAG  1463
451  TrpGluValSerGlnValProLeuAspLeuCysGluValTyrGlyGlyGl  467
1464 TGGGAGGTGAGCCAGGTGCCCCTGGACCTGTGTGAGGTCTATGGCGGGGG  1513
467  yCysHisGlyCysLeuMetSerArgAspProTyrCysGlyTrpAspGlnG  484
1514 CTGCCACGGTTGCCTCATGTCCCGAGACCCCTACTGCGGCTGGGACCAAG  1563
```

```
484  lyArgCysIleSerIleTyrSerSerGluArgSerValLeuGlnSerIle  500
1564 GCCGGCTGCATCCCATCTACAGCTCCGACGGTCAGTGCTGCAATCCATT  1613
501  AsnProAlaGluProHisLysGluCysProAsnProLysProAspLysSerAl  517
1614 AATCCAGCCGAGCCACACAAGGAGTGTCCAAACCCCAAACCAGACAAGGC  1663
517  aProLeuGlnLysValSerLeuAlaProAsnSerArgTyrTyrLeuSerC  534
1664 CCCACTCGCAGAAGGTTTCCCTGGCCCCAAACTCTCGCTACTACCTGAGCT  1713
534  ysProMetGluSerArgHisAlaThrTyrSerArgHisLysLysGluAsn  550
1714 GCCCCATGGAATCCCGCCACGCCACCTACTCACGGCACAAGAAGGAGAAC  1763
551  ValGluGlnSerCysGluProGlyHisGlnSerProAsnCysIleLeuPh  567
1764 GTGGAGCAGAGCTGCGAACCTGGTCACCAGAGCCCCAACTGCATCCTGTT  1813
567  eIleGluAsnLeuThrAlaGlnGlnTyrGlyHisTyrPheCysGluAlaG  584
1814 CATCGAGAACCTCACGGCGCAGCAGTACGGCCACTACTTCTGCGAGGCCC  1863
584  lnGluGlySerTyrPheArgGluAlaGlnGlnHisTrpIleLeuLeuProGlu  600
1864 AGGAGGGCTCCTACTTCCGCGAGGCTCAGCACTGGATCCTGCTGCCCGAG  1913
601  AspGlyIleMetAlaGluAlaHisLeuLeuGlyHisAlaAlaCysAlaLeuAlaI  617
1914 GACGGCATCATGGCCGAGCACCTGCTGGGTCATGCCTGTGCCCTGGCCGC  1963
617  aSerLeuTrpProLeuGlyValLeuProThrLeuThrLeuGlyLeuLeuValH  634
1964 CTCCCTCTGGCCTGGGGGTGCTGCCCACACTCACTCTTGGCTTGCTGGTCC  2013

634  is  634
     ||
2014 AC  2015
```

seq_name: gb_pr4:AF071542

```
seq_documentation_block:
LOCUS       AF071542        2594 bp    mRNA            PRI       01-DEC-1998
DEFINITION  Homo sapiens semaphorin K1 mRNA, complete cds.
ACCESSION   AF071542
NID         g3930578
VERSION     AF071542.1   GI:3930578
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2594)
  AUTHORS   Xu,X., Ng,S., Wu,Z.L., Nguyen,D., Homburger,S., Seidel-Dugan,C.,
            Ebens,A. and Luo,Y.
  TITLE     Human semaphorin K1 is glycosylphosphatidylinositol-linked and
            defines a new subfamily of viral-related semaphorins
  JOURNAL   J. Biol. Chem. 273 (35), 22428-22434 (1998)
  MEDLINE   98380463
REFERENCE   2  (bases 1 to 2594)
  AUTHORS   Xu,X. and Luo,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JUN-1998) Biology, Exelixis Phamaceuticals, Inc., 260
            Littlefield Ave, South San Francisco, CA 94080, USA
REFERENCE   3  (bases 1 to 2594)
  AUTHORS   Xu,X. and Luo,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-NOV-1998) Biology, Exelixis Phamaceuticals, Inc., 260
            Littlefield Ave, South San Francisco, CA 94080, USA
REMARK      Sequence update by submitter
COMMENT     On Dec 1, 1998 this sequence version replaced gi:3493630.
```

```
FEATURES          Location/Qualifiers
  source          1. .2594
                  /organism="Homo sapiens"
                  /db_xref="taxon:9606"
                  /chromosome="15"
                  /map="15q22"
  CDS             1. .2001
                  /note="sema K1; glycosylphosphatidylinositol-linked
                  protein; similar to Alcelaphine herpesvirus 1 semaphorin"
                  /codon_start=1
                  /product="semaphorin K1"
                  /protein_id="AAC80456.1"
                  /db_xref="PID:g3930579"
                  /db_xref="GI:3930579"
                  /translation="MTPPPGRAAPSAPRARVPGPPARLGLPLRLLLLLWAAAASA
                  QGHLRSGPRIFAVWKGHVGQDRVDFGQQEPHTVLFHEPGSSVWVGGRGKVVLFDFPE
                  GKNASVRTVNIGSTKGSCLDKRDCENYITLLERRSEGLLACGTNARHPSCWNLVNGTV
                  VPLGEMRGYAPFSPDENSLVLFEGDEVYSTIRKQEYNGKIPRFRRIRGESELYTSDTV
                  MONPQFIKATIVHQDQAYDDKIYFFRRDNPDKNPEAPLNVSRVAQLCRGDQGGESSL
                  SVSKWNTFLKAMLVCSDAATNKNFNRLQDVFLLPDPSQWRDTRVYGVFSNPWNYSAV
                  CVYSLGDIDKVFRTSSLKGYHSSLPNPRPGKCLPDQQPIPTETFQVADRHPEVAQRVE
                  PMGPLKTPLFHSKYHYQKVAVHRMQASHGETFHVLYLTDRGTIHKVVEPGEQEHSFA
                  FNIMEIQPFRRAAAIQTMSLDAERRKLYVSSQWEVSQVPLDLCEVYGGGCHGCLMSRD
                  PYCGWDQGRCISYSSERSVLQSINPAEPHKECPNPKPDKAPLQKVSLAPNSRYYLSC
                  PMESRHATYSWRRKENVEQSCEPGHQSPNCILFIENLTAQQYHYFCEAQEGSYFREA
                  QHWQLLPDGIMAEHLLGHACALAASLNLGVLPTLTGLLVH"

BASE COUNT       514 a       814 c       756 g       510 t
ORIGIN

alignment_scores:
   Quality:    3443.00                Length:     634
   Ratio:         5.431               Gaps:         0
Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
US-09-041-236-2 x AF071542

Align seg 1/1  to: AF071542  from: 1  to: 2594

  1 LeuLeuLeuLeuTrpAlaAlaAlaAlaSerAlaGlnGlyHisLeuAr 17
    |||||||||||||||||||||||||||||||||||||||||||||||
 97 CTGCTGCTGCTCTGGGCGGCGGCCGCGGCTCCGCCCAGGGCCACCTAAG 146

 17 gSerGlyProArgIlePheAlaValTrpLysGlyHisValGlyGlnAspA 34
    ||||||||||||||||||||||||||||||||||||||||||||||||||
147 GAGCGGACCCCGCATCTTCGCCGTCTGGAAAGGCCATGTAGGGCAGGACC 196

 34 rgValAspPheGlyGlnThrGluProHisThrValLeuPheHisGluPro 50
    ||||||||||||||||||||||||||||||||||||||||||||||||||
197 GGGTGGACTTTGGCCAGACTGAGCCGCACCGTGCTTTTCCACGAGCCA 246

 51 GlySerSerValTrpValGlyValArgGlyLysValValLeuPheAs 67
    |||||||||||||||||||||||||||||||||||||||||||||||
247 GGCAGCTCCTCGTGTGGGTGGGAGGACGTCTACCCTCTTGA 296

 67 pPheProGluGlyLysAsnAlaSerValArgThrValAsnIleGlySerT 84
    ||||||||||||||||||||||||||||||||||||||||||||||||||
297 CTTCCCGGAGGGCAAGAACGCCATCTGTGCCACGGTGAATATCGGCTCCA 346

 84 hrLysGlySerCysLeuAspLysArgAspCysGluAsnTyrIleThrLeu 100
    ||||||||||||||||||||||||||||||||||||||||||||||||||
347 CAAAGGGGTCCTGTCTGGATAAGCGGGACTGCGAGAACTACATCACTCTC 396

101 LeuGluArgArgSerGluGlyLeuLeuAlaCysGlyThrAsnAlaArgHi 117
    ||||||||||||||||||||||||||||||||||||||||||||||||||
397 CTGGAGCGGCGGAGTGAGGGGCTGCTGGCGTGCGGCACCAACGCCCGGCA 446

117 sProSerCysTrpAsnLeuValAsnGlyThrValProLeuGlyGluM 134
    |||||||||||||||||||||||||||||||||||||||||||||||
447 CCCCAGCTGCTGGAACCTGGTGAACGGCACTGTGCCACTTGGCGAGA 496

134 etArgGlyTyrAlaProPheSerProAspGluAsnSerLeuValLeuPhe 150
```

```
                           |||||||||||||||||||||||||||||||||||||||||||||||||||||
                      497 TGAGAGGCTACGCCCCTTCAGCCCGGACGAGAACTCCCTGTTCGTTT 546

151 GluGlyAspGluValTyrSerThrIleArgLysGlnGluTyrAsnGlyLy 167
    |||||||||||||||||||||||||||||||||||||||||||||||||
547 GAAGGGGACAGAGGTGTATTCCACCATCCGGAAGCAGGAATACAATGGGAA 596

167 sIleProArgPheArgArgIleArgGlySerGluLeuTyrThrSerA 184
    |||||||||||||||||||||||||||||||||||||||||||||||
597 CATCCCTCGGTTCCGCCGCATCCGGGGCGAGAGTGAGCTGTACACCAGTG 646

184 spThrValMetGlnAsnProGlnPheIleLysAlaThrIleValHisGln 200
    ||||||||||||||||||||||||||||||||||||||||||||||||||
647 ATACTGTCATGCAGAACCCACAGTTCATCAAAGCCACCATCGTGCACCAA 696

201 AspGlnAlaTyrAspAspLysIleTyrTyrPhePheArgArgAspAsnPr 217
    ||||||||||||||||||||||||||||||||||||||||||||||||||
597 GACCAGGCTTACGATGACAAGATCTACTACTTCTTCCGAGAGGACAATCC 746

217 oAspLysAsnProGluAlaProLeuAsnValSerArgValAlaGlnLeuC 234
    ||||||||||||||||||||||||||||||||||||||||||||||||||
747 TGACAAGAATCCTGAGGCTCCTCAATGTGTCCCGTGTGGCCCAGTTGT 796

234 ysArgGlyAspGlnGlyGlyGluSerSerLeuSerValSerLysTrpAsn 250
    ||||||||||||||||||||||||||||||||||||||||||||||||||
797 GCAGGGGGACCAGGGTGGGGAAAGTCACTGTCAGTCTCCAAGTGGAAC 846

251 ThrPheLeuLysAlaMetLeuValCysSerAspAlaAlaThrAsnLysAs 267
    ||||||||||||||||||||||||||||||||||||||||||||||||||
847 ACTTTTCTGAAGCCATGCTGGTATGCAGTGACGTGATGCTGCCACCAACAAGAA 896

267 nPheAsnArgLeuGlnAspValPheLeuLeuProAspProSerGlnIrnT 284
    |||||||||||||||||||||||||||||||||||||||||||||||||
897 CTTCAACAGGCTGCAAGACGTCTTCCTGCTCCCTGACCCCAGCCAGCCAGT 946

284 rpArgAspThrArgValTyrGlyValPheSerAsnProTrpAsnTyrSer 300
    ||||||||||||||||||||||||||||||||||||||||||||||||||
947 GGAGGGACACCAGGGTCTATGGTGTTTTCTCAACCCCTGGAACTACTCA 996

301 AlaValCysValTyrSerLeuGlyAspIleAspLysValPheArgThrSe 317
    ||||||||||||||||||||||||||||||||||||||||||||||||||
997 GCCGTCTGTGTGTATTCCCCGGTGACATTGACAAGGTCTTCCGTACCTC 1046

317 rSerLeuLysGlyTyrHisSerSerLeuProAsnProArgProGlyLysC 334
    ||||||||||||||||||||||||||||||||||||||||||||||||||
1047 CTCACTCAAGGGCTACCACTCAAGCCTTCCCAACCCGCGGCCTGGCAAGT 1096

334 ysLeuProAspGlnGlnProIleProThrGluThrPheGlnValAlaAsp 350
    ||||||||||||||||||||||||||||||||||||||||||||||||||
1097 GCCTCCCAGACCAGCAGCCGGATACCCCACAGAGACCTTCAGGTGGCTGAC 1146

351 ArgHisProGluValAlaGlnArgValGluProMetGlyProLeuLysTh 367
    ||||||||||||||||||||||||||||||||||||||||||||||||||
1147 CGTCACCCAGAGGTGGCGCAGAGGGTGGAGCCCATGGGGCCTCTGAAGAC 1196

367 rProLeuPheHisSerLysTyrHisTyrGlnLysValAlaValHisArgM 384
    ||||||||||||||||||||||||||||||||||||||||||||||||||
1197 GCCATTGTTCCACTCTAAATACCACTACCAGAAAGTGGCCGTCCACCGCA 1246

384 etGlnAlaSerHisGlyGluThrPheHisValLeuTyrLeuThrThrAsp 400
    ||||||||||||||||||||||||||||||||||||||||||||||||||
1247 TGCAAGGCCAGCCACGGGGAGACCTTTCATGTGCTTTACCTAACTACAGAC 1296

401 ArgGlyThrIleHisLysValValGluProGlyGluGlnHisSerPh 417
    |||||||||||||||||||||||||||||||||||||||||||||||
1347 AGGGGCACTATCCACAAGGTGTGGAACCGGGGAGCAGGACCACACGTT 1346

417 eAlaPheAsnIleMetGluIleGlnProPheArgArgAlaAlaAlaIleG 434
    ||||||||||||||||||||||||||||||||||||||||||||||||||
347 CGCCTTCAACATCATGGAGATCCAGCCCTTCCGCCGCGGCTGCCATCC 1396

434 lnThrMetSerLeuAspAlaGluArgArgLysLeuTyrValSerSerGln 450
    ||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
1397 AGACCATCTCGCTGGATCTGAGCGGAGGAAGCTGTATGTGAGCTCCCAG 1446
 451 TrpGluValSerGlnValProLeuAspLysCysGlyValTyrGlyGlyGl 467
     |||||||||||||||||||||||||||||||||||||||||||||
1447 TGGGAGGTGAGCCAGGTGCCCCTGGACCTGTGTGAGGCTATGGCGGGGG 1496

 467 yCysHisGlyCysLeuMetSerArgAspPheTyrCysGlyTrpAspGlnG 484
     |||||||||||||||||||||||||||||||||||||||||||||
1497 CTGCCACGGTGCCTCATGTCCCGAGACCCCTACTGCGGCTGGGACCAAG 1546

 484 lyArgCysIleSerIleTyrSerSerGluArgSerValLeuGlnSerIle 500
     |||||||||||||||||||||||||||||||||||||||||||||
1547 GCGCGTGCATCTCCATCTACAGCTCCGAACGGTCAGTGCTGCAAATCCATT 1596

 501 AsnProAlaGluProHisLysGlyCysProAsnProLysProAspLysSerAl 517
     |||||||||||||||||||||||||||||||||||||||||||||
1597 AATCCAGCCGAGCCACACAAGGAGTGTCCAACCCCAAACCAGACAAGGC 1646

 517 aProLeuGlnLysValSerLeuAlaProAsnSerArgTyrTyrLeuSerC 534
     |||||||||||||||||||||||||||||||||||||||||||||
1647 CCCACTGCAGAAGGTTTCCCTGGCCCCAAACTCTCGCTACTACCTGAGCT 1696

 534 ysProMetGluSerArgHisAlaThrTyrSerTrpArgHisGluGluAsn 550
     |||||||||||||||||||||||||||||||||||||||||||||
1697 GCCCCATGGAATCCCGCCACGCCACCTACTCATGGCGCCACGAGGAGAAC 1746

 551 ValGlyGlnSerCysGluProGlyHisGlnSerProAsnCysIleLeuPh 567
     |||||||||||||||||||||||||||||||||||||||||||||
1747 GTGGGACAGAGCTGCGAACCTGGTCACCAGAGCCCCAACTGCATCCTGTT 1796

 567 eIleGluAsnLeuThrAlaGlnGlnTyrGlyHisTyrPheCysGluAlaG 584
     |||||||||||||||||||||||||||||||||||||||||||||
1797 CATCGAGAACCTCACGGCGCAGCAGTACGGCCACTACTTCTGCGAGGCCC 1846

 584 lnGluGlySerTyrPheArgGluAlaGlnGlnHisTrpGlnLeuProGlu 600
     |||||||||||||||||||||||||||||||||||||||||||||
1847 AGGAGGGCTCCTACTTCCGCGAGGCCCAGCAGCACTGGCAGCTGCCCGAG 1896

 601 AspGlyIleMetAlaGluHisLeuLeuLeuGlyHisAlaCysAlaLeuAlaAl 617
     |||||||||||||||||||||||||||||||||||||||||||||
1897 GACGGCATCATGGCCGAGCACCTGCTGGGTCATGCCTGCGCCCTGGCCGC 1946

 617 aSerLeuTrpLeuGlyValValLeuProThrLeuThrLeuGlyLeuLeuValH 634
     |||||||||||||||||||||||||||||||||||||||||||||
1947 CTCCCTCTGGCTGGGGGTGGTGCTGCCCCACTCACTCTTGGCTTGCTGGTCC 1996

634 1s 634
    ||
1997 AC 1998
```

seq_name: gb_ro:AB017532

seq_documentation_block:
```
LOCUS       AB017532      2608 bp    mRNA        ROD        09-MAR-1999
DEFINITION  Mus musculus mRNA for msemK1p, complete cds.
ACCESSION   AB017532
NID         g4519587
VERSION     AB017532.1  GI:4519587
KEYWORDS    msemK1p.
SOURCE      Mus musculus postnatal brain cDNA to mRNA.
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (sites)
AUTHORS     Sato,Y. and Takahashi,H.
TITLE       Molecular cloning and expression of novel Semaphorin family of
            molecule
JOURNAL     Unpublished (1998)
REFERENCE   2  (bases 1 to 2608)
AUTHORS     Takahashi,H. and Sato,Y.
TITLE       Direct Submission
JOURNAL     Submitted (08-SEP-1998) to the DDBJ/EMBL/GenBank databases. Hiroshi
            Takahashi, Mitsubishi Kasei Institute of Life Sciences,
```

Neurobiology Lab; 11 Minamiooya, Machida, Tokyo 194, Japan
(E-mail:hiroshi@libra.ls.m-kagaku.co.jp, Tel:81-427-24-6211,
Fax:81-427-24-6314)

```
FEATURES             Location/Qualifiers
source               1..2608
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /dev_stage="postnatal"
                     /tissue_type="brain"
gene                 6..2000
                     /gene="msemK1"
CDS                  6..2000
                     /gene="msemK1"
                     /codon_start=1
                     /product="msemK1p"
                     /protein_id="BAA75665.1"
                     /db_xref="PID:d1039405"
                     /db_xref="PID:g4519588"
                     /db_xref="GI:4519588"
                     /translation="MTPPPGRAAPSAPRARVLSLPARFGLPLRLRLLLVFWVAAASA
                     QGHSRSGPRISAVWKGQDHVDFSQPEPHTVLFHEPGSFSVWVGGRGKVYHFNFPEGKN
                     ASVRTVNIGSTKGSCQDKQDCGNYITLLERRGNGLLVCGTNARKPSCWNLVNDSVVMS
                     LGEMKGYAPFSPDENSLVLFEGDEVYSTIRKQEYNGKIPRFRRIRGESELYTSDTVMQ
                     NPQFIKATIVHQDQAYDDKIYYFFREDNPDKNPEAPLNVSRVAQLCRGDQGGESSLSV
                     SKWNTFLKAMLVCSDAATNRNFNRLQDVFLLPDPSGQWRDRVYGVFSNPWNYSAVCV
                     YSLGDIDRVFRTSSLKGYHMGLPNPRPGMCLPKKQPIPTETFQVADSHPEVAQRVEPM
                     GPLKPLFHSKYHYQKVVVHRMQASNGETFHVLYLTTDRGTIHKVVESGDQDHSFVFN
                     IMEIQPFHRAAAIQRISLDADRRKLYVTSQWEVSQVPLDMCEVYSGGCHGCLMSRDFY
                     CGWDQDRCVSYSSQRSVLQSINPAEPHRECPNPKPDEAPLQKVSLARNSRYYLTCPM
                     ESRHATYLWRHEENVEQSCEPGHQSPSCCILFIENLTARQYGHYRCEAQEGSYLREAQH
                     WELLPEDRALAEQLMGHARALAASFWLGVLPFTLILGLLVH"

BASE COUNT    554 a    769 c    727 g    558 t
ORIGIN
```

alignment_scores:
```
         Quality: 3072.00          Length:       635
           Ratio: 5.053            Gaps:          2
Percent Similarity: 95.748    Percent Identity: 89.134
```

alignment_block:
US-09-041-236-2 x AB017532

Align seg 1/1 to: AB017532 from: 1 to: 2608

```
  1 LeuLeuLeuLeuTrPAlaAlaAlaAlaSerAlaGlnGlyHisLeuAr 17
    ||||||||||||::||||:|||||||||||||||||||||||||| ||
102 CTTCTGCTGCTGGTGTTCTGGCCAGCCGCCGCCGCCCTCCGCCCCAAGGCCACTCGAG 151

 17 gSerGlyProArgIlePheAlaValTrpLysGlyHisValGlyGlnAspA 34
    ||||||||||| ||:||| ||||||::|| ||||||:||::::|| ||||
152 GAGCGGACCCCGCATCTCCGCCGTCTGGAAA........GGCCAGGACC 192

 34 rgValAspPheGlyGlnThrGluProHisThrValLeuPheHisGluPro 50
    :|||||||||:|||| ||||||||||||||||||||||||||||||||||
193 ATGTGGGACTTAGCCAGCCTGAGCCACCACACCGTGCTTTTCCATGAGCCG 242

 51 GlySerSerSerValTrpValGlyValGlyArgGlyLysValTyrLeuPheAs 67
    :|||||||||:||| |||||||||||||||||||:|||||||| || |||::
243 GGCAGCTCTCTGTCTGGGTGGGTGGTGGACGTGGCAAGGTCTACCACTTCAA 292

 67 pPheProGluGlyLysAsnAlaSerValAlaServalAlaAsnIleGlySerT 84
    :|||||||||||||||||:|| |::||||::::|||::|||||||||||| ||
293 CTTCCCCGAGGGCAAGAATGCCTCTGTGCCACGGTGAACATCGGCCTCCA 342

 84 hrLysGlySerCysLeuAspLysArgAspCysGluAsnTyrIleThrLeu 100
    |||||||||||| |:|:||| || || || ||:||||||||||||||
343 CAAAGGGTCCTGCAGGACAAACAGGACTGTGGGAATTACATCACTCTT 392

101 LeuGluArgArgSerGluGlyLeuLeuAlaCysGlyThrAsnAlaArgH 117
    ||||||||:|| ::::|||::|||||||||||:|||||||||||||||
393 CTAGAAAGGCGGGGTAATGGCTGCGTCTGTGGCACCAATGCCCGGAA 442
```

```
117 sProSerCysTrpAsnLeuAsnGlyThrValVal...ProLeuGlyG 133
    |||||||||||||||||||||      ::|||||||||       ||||||
443 GCCCAGCTGCTGGAACTTGGTGACAGTGTGGTGATGTCACTTGGTG 492

133 luMetArgGlyTyrAlaProPheSerProAspGluAsnSerLeuValLeu 149
    |||||::|||||||||||||||||||||||||||||||||||||||||||
493 AGATGAAAGGCTATGCCCCCTCAGCCCCGGATGAGAACTCCCTGGTTCTG 542

150 PheGluGlyAspGluValTyrSerThrIleArgLysGlnGluTyrAsnGl 166
    |||||||||||||||||||||||||||||||||||||||||:||
543 TTTGAAGGAGATGAAGTGTACTCTACCATCCGGAAGCAGGAATACAACGG 592

166 yLysIleProArgPheArgArgIleArgGlyGluSerGluLeuTyrThrS 183
    ||||||||||||||||||||||||||||||||||||||||||
593 GAAGATCCCTCGGTTTCGACGCATTCGGGGCGAGAGTGAACTGTACACAA 642

183 erAspThrValMetGlnAsnProGlnPheIleLysValAlaThrIleValHis 199
    |||||||||||||||||||||||||||||||||||||||||||
643 GTGATACAGTCATGCAGAACCCACAGTTCATCAAGGCCACCATTGTGCAC 692

200 GlnAspGlnAlaTyrAspAspLysIleTyrTyrPhePheArgGluAspAs 216
    ||||||||||||||||||||:|||||||||||||||||||||||
693 CAAGACCAAGCCTATGATGATAAGATCTACTACTTCTTCCGAGAAGACAA 742

216 nProAspLysAsnProGluAlaProLeuAsnValSerArgValAlaGlnL 233
    |||||||||||||||||||||||||||||||:||:||||||
743 CCCTGACAAGAACCCCGAGGCTCCTCTCAATGTGTCCCGAGTAGCCCAGT 792

233 euCysArgGlyAspGlnGlyGlyGluSerLeuSerValSerLysTrp 249
    |||||||||||||||||||||||||||||||||||||||
793 TGTGCCGAGGGGGACCAGGGGGTGGAGAGTTCGTTGCTGTCTCCAAGTGG 842

250 AsnThrPheLeuLysAlaMetLeuValCysSerAspAlaAlaThrAsnLy 266
    |||||||||||||||||||||||||||||||||::
843 AACACCTTCCTGAAAGCCATGTTGGTCTGCAGCGATGCAGCCACCAACAG 892

266 sAsnPheAsnArgLeuGlnAsnProValPheLeuLeuProAspProSerGlyG 283
    :||||||||||||||||||||||||||||||||||||
893 GAACTTCAATCGGCTGCAAGATGTCTTCCTGCTCCCCGACCCCAGTGGCC 942

283 lnTrpArgAspThrArgValTyrGlyValPheSerAsnProTrpAsnTyr 299
    |||||||||||||||||||||||||||||||||||||||
943 AGTGGAGAGATACCAGGGGTCTATGGCGTTTTCTCCAACCCCTGGAACTAC 992

300 SerAlaValCysValTyrSerLeuGlyAspPheIleAspLysValPheArgTh 316
    |||||||||||||||||||||||||||||||
993 TCAGCTGTCTGCGTGTATTCGCTTGGTGACATTGACAGAGTCTTCCGTAC 1042

316 rSerSerLeuLysGlyTyrHisSerSerLeuSerProAsnProArgProGlyL 333
    ||||||||||||::|||||||||||||||
1043 CTCATCGCTCAAAGGCTACCACATGGGCCTTCCCAACCCTCGACCTGGCA 1092

333 ysCysLeuProAspGlnProIleProThrGluThrPheGluValAlaAla 349
    |||||  :::|||||||||||||||||||
1093 TGTGCCTCCCAAAAAGCAGCCCATGGGCCATACACCCCACAGAAACCTTCCAGGTAGCT 1142

350 AspArgHisProGluValAlaGlnValGlnArgValGluProMetGlyTyrProLeuLy 366
    ||||||||||||||||||||||||||||||||
1143 GATAGTCACCCAGAGGTGGCTCAGAGGGTCGAACCCATGGGGCCACTGAA 1192

366 sThrProLeuPheHisSerLysTyrHisSerTyrGlnLysValAlaValHisAla 383
    ||||||||||||||||||||||||::|||||
1193 GACACCATTGTTCCATTCTAAGTACCATTACCAGAAAGTGGTCGTTCACC 1242

383 rgMetGlnAlaSerHisGlyValGluThrPheHisValLeuTyrLeuThrThr 399
    |||::||||||||||||||||||||||
1243 GCATCCAAGCCAGCCAGCAATGGAGAGACCTTCCATGTGCTTTATCTAACCACA 1292

400 AspArgGlyThrIleHisLysValValGluProGlyGluGlnHisSe 416
    |||||  :::||||||||||||||||||||
1293 GACAGGGGCACCATTCACAAGGTGGTGGAATCAGGGACCAGGACCATAG 1342

416 rPheAlaPheAsnIleMetGluIleGlnProPheArgArgAlaAlaAlaAlaI 433


1343 CTTTGTCTTCAATATCATGGAGATCCAGCCCTTCACGTGCAGCGGCCA 1392

433 leGlnThrMetSerLeuAspAlaArgAlaArgLysLeuTyrValSerValSerSer 449
    |||||||||||||||||||||||::||||
1393 TCCAGGCTATATCATTGGATCTGACGCCGGCGAAGCTCTATGTGACCTCC 1442

450 GlnTrpGluValSerGlnValProLeuAspLeuCysGluLeuValTyrGlyGl 466
    |||||||||||:|||||::||||||||||||:::
1443 CAGTGGGAAGTGAGCCAGGTACCCCCTGGACATGTGTGAGGTCTACAGCGG 1492

466 yGlyCysHisGlyCysLeuMetSerArgAspProTyrCysGlyTrpAspG 483
    |||||||||||||||||||||||||||||||||||||
1493 GGGCTGCCATGGCTGCCTCATGTCCCGAGACCCCTACTGTGGCTGGGACC 1542

483 lnGlyArgCysIleSerIleTyrSerSerGluArgSerValLeuGlnSer 499
    |  ||||||||||||||||||::||||||||||
1543 AGGACCGCTGCGTGTCTATCTACAGCTCCACGATCAGTGCTGCAGTCC 1592

500 IleAsnProAlaGluProHisLysGluCysProAsnProLysProAspLy 516
    ||||||||:||:|||::|||||::|||||
1593 ATTAATCCAGGGAGCCACACAGAGTGTCCCAACCCTAAACCAGACGA 1642

516 sAlaProLeuGlnLysValSerLeuAlaProAsnSerArgArgTyrTyrLeuS 533
    |||||||||||||||||||||||:||:|||||||
1643 GGCTCCATTGCAGAAGGTTCCCTGGCTCGGAACTCCGATACTACCTGA 1692

533 erCysProMetGluSerArgHisAlaThrTyrSerTrpArgHisLysGlu 549
    ::|||||||||::||||||||||||||:|||||
1693 CCTGCCCCATGGAGTCCCGCCACGCCACTTACTTATGGCGCCATGAGGAG 1742

550 AsnValGluGlnSerCysGluProGlyHisGlnSerProAsnCysIleIle 566
    |||||||||||||||||||||||||||||||:|||||
1743 AATGTGGAACAGAGCTGTGAGCCCGGCCACCAAAGCCCTAGCTGCATCCT 1792

566 uPheIleGluAsnLeuThrHrAlaGlnGlnTyrGlyHisTyrPheCysGluA 583
    ||||||||||||::|||||||||::|||||
1793 GTTCATTGAGAACCTCACGGCCCGTCAGTATGGCCACTACCCCTGCGAGG 1842

583 laGlnGluGlySerTyrPheArgGluAlaGlyGlnHisTrpGlnLeuLeuPro 599
    |||||||||||||::|||||||||||||||
1843 CCCAAGAGGGCTCCTACCTTCCGTGAGGCTCAACACTGGCAGCTGCTGCCA 1892

600 GluAspGlyIleMetAlaGluAlaGluHisLeuLeuGlyHisAlaCysAlaLeuAl 616
    ::|||||||||||||||:||||||
1893 GAGGACGGAGGCACTGGCCGAACAACTAATGGGCCATGCCCGGCCCCTGGC 1942

616 aAlaSerLeuTrpLeuGlyValLeuProThrLeuThrLeuGlyLeuLeuV 633
    ||||||||||||||||||||||||||||
1943 CGCCTCCTTGTGGCTCGGGGTCCTGCCCACACTCATACTTGGTCTGCTGG 1992

633 alHis 634
    |||||
1993 TTCAC 1997
```

REFERENCE   2 (bases 1 to 2538)
AUTHORS   Mine,T., Harada,K., Matsumoto,T., Yamana,H., Itoh,K. and Yamada,A.
TITLE   Direct Submission
JOURNAL   Submitted (10-AUG-1999) Immunology, Kurume University School of
          Medicine, Asahi-machi 67, Kurume, Fukuoka 830-0011, Japan
FEATURES   Location/Qualifiers
   source   1..2538
            /organism="Mus musculus"
            /strain="C57BL/6J"
            /db_xref="taxon:10090"
            /chromosome="9"
            /map="9A3.3-B"
   gene   1..2538
            /gene="Cdw108"
   CDS   6..2000
            /gene="Cdw108"
            /note="similar to Homo sapiens CDw108"
            /codon_start=1
            /product="GPI-anchored membrane protein CDw108 precursor"
            /protein_id="AAD53118.1"
            /db_xref="PID:g5823361"
            /db_xref="GI:5823361"
            /translation="MTPPPPGRAAPSAPRARVLSLPARFGLPLRLRLLLVFWVAAASA
            QGHSRSGPRISAVWKGQDHVDFSQPEPHTVLFHEPGSFSVWVGGRGKVVHFNFPEGKN
            ASVRTVNIGSTKGSCQDKQDGNYITLLERRGNGLLVCGTNARKPSCWNLVNDSVVMS
            LGEMGYAPFSPDENSLVLFEGDEYYSTIRKQEYNGKIPRFRRIRGESELYTSDTVMQ
            NPQFIKATIVHQDQAYDDKIYFFREDNPDKNPEAPLNVSRVAQLCRGDQGGESSLSV
            SKWNFLKAMLVCSDAATNRNFNRLIQDVFLLPDFSGQWRDTRVIGVFSNPWNYSAVCV
            YSLGDIDRVFRTSSLKGYHMGLPNPRPGMCLPKKQPIPTETFQVADSHPEVAQRVEPM
            GPLKTPLFHSKYHYQKVVTHRMQASNGEITFHVLLTTDRGTIHKVVESGDQDHSFVEN
            IMEIQPFHRAAAIQAISLDADRRKLYVTSQWEVSQVPLDMCEVYSGGCHGCLMSRDPY
            CGWDQDRCVSIYSSQRSVLQSINPAEPHRECPNPKPDEAPLQKVSLARNSRYYLTCPM
            ESRHATYLWRHEENVEQSCEPGHQSPSCLFIENLTARQYGHYRCEAQEGSYLREAQH
            WELLPEDRALAEQLMGHARALAASFWLGVLPTLIGLLVH"
BASE COUNT   538 a   757 c   713 g   530 t
ORIGIN

alignment_scores:
      Quality: 3072.00   Length: 635
      Ratio: 5.053   Gaps: 2
Percent Similarity: 95.748   Percent Identity: 89.134

alignment_block:
US-09-041-236-2 x AF176670   .:

Align seg 1/1   to: AF176670   from: 1   to: 2538

   1 LeuLeuLeuLeuTrpAlaAlaAlaAlaAlaSerAlaGlnGlyHisLeuAr 17
     |||||||||||||::|||| |||||||||||||||||||| ||:||||
 102 CTTCTGCTGGTTCTGGGTGGCCGCCGCGCCTCCGCCCAAGGCCACTCGAG 151

  17 gSerGlyProArgIlePheAlaValTrpLysValGlyHisGlyHisAspA 34
     ||||||||||||||||||||||| |||||||||| ||||||||
 152 GAGCGGGACCCCGCATCTCCGCGTCTGGAAA.........GGGCAGGACC 192

  34 rgValAspPheGlyGlnThrGluProHisThrValLeuPheHisGluPro 50
     ||:|||||||||||::|||||||||||||||||||||||||||
 193 ATGTGGACTTTAGCCAGCCTGAGCCCAGCCACACCCGTGCTTTCCATGAGCCCG 242

  51 GlySerSerValTrpValGlyGlyArgGlyLysValTyrLeuPheAs 67
     |||||||||||||||| ||||||||||||||||||:|||:|||
 243 GGCAGCTTCTCTGTCTGGGTGGGTGGACGTGGCAAGGTCTACACCTTCAA 292

  67 pPheProGluGlyLysAsnAlaSerValArgThrValAsnIleGlySerT 84
     ||:||||||||||||||||||||||:||::|||||||||||||||||
 293 CTTTCCAGAGGGCAAGAATGCCTCTGTGCGCACGGTGAACATCGGCTCCA 342

  84 hrLysGlySerCysLeuAspLysArgAspCysLeuAsnTyrIleThrLeu 100
     ||||||||||||||||||:|||::||||:|||||||||||||||||
 343 CAAAGGGGTCCTGCAGGACACAAACAGGACTGTGGAATTACATCACTCTT 392

 101 LeuGluArgArgSerGluGlyLeuLeuValCysGlyThrAsnAlaArgHi 117
     |||||||||||||||||||||||||||::||||:|||||||||||||

       |||||||||||||||||||||||||:::||||||||||||||||||||||
   393 CTAGAAAGGCGGGGTAATGGGCTGCTGTGTCTGTGGCCACCAATGCCCGAA 442

   117 sProSerCysTrpAsnLeuValAsnGlyThrValVal...ProLeuGlyG 133
       ||||||||| |||||||||||||||||| :::||||| |||||:|
   443 GCCCAGCTGCTGGTGAACTTGGTGAATGACAGTGGTGATGTCACTTGGTG 492

   133 luMetArgIleTyrAlaProPheSerProAlaGlnSerLeuValLeu 149
       ||||:::||| ||||||||||||||| |||||||||||||||||||||||
   493 AGATGAAAGCTATGCCCCCTCAGCCCGGATGAGAACTCCCTGGTTCTG 542

   150 PheGluGlyAspGluValTyrSerThrIleArgGlnGluTyrAsnGl 166
       ||||||||||||||||||||||||||||||||||||||||||||||||
   543 TTTGAAGGAGATGAAGTGTACTCTACCATCCGACAGGAATACAACGG 592

   166 yLysIleProArgPheArgArgIleArgGlySerGluLeuTyrThrS 183
       ||||||||||||||||||||||||||||||||||||||||||||:||:
   593 GAAGATCCCTCGTTTCGACGGTTCGGGGCGAGAGTGAACTGTACACAA 642

   183 erAspThrValMetGlnAsnProGlnPheIleLysAlaThrIleValHis 199
       |||||||||| |||||||||||||||||||||||||||||||||||||||
   643 GTGATACAGTCATGCAGAACCCACAGTTCATCAAGGCCACCATTGTCAC 692

   200 GlnAspGlnAlaTyrAspAspLysIleTyrPhePheArgGluAspAs 216
       ||||||||| |||||||:|||||||||||||:|||||||||||||||||
   693 CAAGACCAAGCCTATGATGATAAGATCTACTACTTCTTCCGAGAAGACAA 742

   216 nProAspLysAsnProGluAlaProLeuAsnValSerArgValAlaGlnL 233
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   743 CCCTGACAAGAACCCCGAGGCTCCTCTCAATGTCCCAGTAGCCCAGT 792

   233 euCysArgGlyAspGlnGlyGlyGluSerSerLeuSerValSerTrp 249
       |||||||||||||||||||||||||||||||||||||||||||||::
   793 TGTGCAGGGGGGACCAGGGTGAGAGTTCGTTGTCTGTCTCCAAGTGG 842

   250 AsnThrPheLeuLysAlaMetLeuValCysSerAspAlaAlaThrAsnLy 266
       :::|||||||||||||||||||||||||||||||||||||||||||||||
   843 AACACCTTCCTGAAAGCCATGTTGGTCTGCAGCGATGCAGCCACCAACAG 892

   266 sAsnPheAsnArgLeuGlnAspValPheLeuLeuProAspProSerGlyG 283
       :||||||||||||||||||||||||||||||||||||||||||||||||
   893 GAACTTCAATCGGCTGCAAGATGTCTTCCTGCTCCCTGACCCCAGTGGCC 942

   283 lnTrpArgAspThrArgValIleGlyValPheSerAsnProTrpAsnTyr 299
       ||||||||||||||||||||||||||||||||||||||||||||||
   943 AGTGGAGAGATACCAGGGTCATAGGCCGTTTCTCCAACCCCTGGAACTAC 992

   300 SerAlaValCysValTyrSerLeuGlyAspIleAspArgValPheArgTh 316
       ||:|||||||||||||||||||||||||:::||||||||||||||||||
   993 TCAGCTGTCTGCGTGTATTCGCTGGGTGACATTGACAGAGTCTTCCGGTAC 1042

   316 rSerSerLeuLysGlyTyrHisSerLeuProAsnProArgProGlyL 333
       |||||||||||||||||||||||||||||||||||||||||||||||
  1043 CTCATCGCTCAAAGGCTACCACATGGGCCTTCCCAACCCTCGACCTGGCA 1092

   333 ysCysLeuProLysGlnProIleProThrGluThrPheGlnValAla 349
       |||||||||| :::|||||||||||||||||||||||||||||||||
  1093 TGTGCCTTCCCAAAAAAGCAGCCCATACCCACAGAAAACCTTCAGGTAGCT 1142

   350 AspArgHisProGluAlaGlnArgValGluProMetGlyProLeuLy 366
       ||||||||||||||||||||||||||||||||||||||||||||||||
  1143 GATAGTCACCCAGAGGTGGCTCAGAGGGTGGAACCTATGGGGCCACTGAA 1192

   366 sThrProLeuPheHisSerLysTyrHisTyrGlnLysValValThrHisA 383
       ||||||||||||||||||||||||||||||||||||||||||||||||||
  1193 GACACCATTGTTCCATTCTAAGTACCATTACCAGAAAGTGGTCGTTCACC 1242

   383 rgMetGlnAlaSerHisGlyGluThrPheHisValLeuTyrLeuThrThr 399
       |||||||:||||:|||||||||||||||||||||||||||||||||||||
  1243 GCATGCAAGCCAGCAATGAGGAGAACCTTCCATGTGCTTTATCTAACCACA 1292

   400 AspArgGlyThrIleHisLysValValGluSerGlyGlnGlnHisSe 416
       |||||||||||||||:::||||||||||||||||||||||||||

TITLE   Ensser,A.
New eukaryotic semaphorins with close homology to semaphorins of
DNA viruses
JOURNAL   Genomics 51 (3), 340-350 (1998)
MEDLINE   98389619
REFERENCE 2 (bases 1 to 8861)
AUTHORS   Ensser,A., Liehr,T., Lange,C., Goen,M., Gebhart,E. and
Fleckenstein,B.
TITLE   Direct Submission
JOURNAL   Submitted (22-OCT-1997) Institut fuer Klinische und Molekulare
Virologie, Friedrich Alexander Universitaet, Schlossgarten 4,
Erlangen, Bavaria 91054, Germany

FEATURES   Location/Qualifiers
source   1..8861
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="15p22.3-p23"
/chromosome="15"
mRNA   join(<1..114,605..646,944..1036,1238..1322,1443..1553,
2176..2315,2907..3091,3945..4053,4146..4344,6876..7003,
7178..7332,7533..7594,7904..>8265)
/gene="SEMAL"
/product="Semaphorin L"
/evidence=experimental
gene   <1..>8265
/gene="SEMAL"
CDS   join(<1..114,605..646,944..1036,1238..1322,1443..1553,
2176..2315,2907..3091,3945..4053,4146..4344,6876..7003,
7178..7332,7533..7594,7904..8265)
/gene="SEMAL"
/note="sequence starts with the putative second exon of
SEMAL; H-SEMA-L"
/codon_start=1
/product="semaphorin L"
/protein_id="AAC34741.1"
/db_xref="PID:g3551779"
/db_xref="GI:3551779"
/translation="EPHTVLFHEPGSSSVWVGGRGKVYLFDFPEGKNASVRTVNIGST
KGSCLDKRDCENYITLLERSEGLLACGTNARHPSCWNLVNGTVVPLGEMRGYAPFSP
DENSLVLFEGDEVYSTIRKQEYNGKIPRFRIGESELYTSDTVMQNPQFIKATIVHQ
DQAYDDKIYYFFREDNPDKNPEAPLNVSRVAQLCRDQGGESSLSVSKWNTFLKAMLV
CSDAATNKNFNRLQDVFLLPDPSGQWRDTRVYGVFSNPWNYSAVCVYSLGDIDKVFRT
SSLKGYHSSLPNPRPGKCLPDQQPIPTEFQVARDRHPEVAQRVEPMGPLKTPLFHSKY
HYQKVAVHRMQASHGETFHVLYLTTDRGTIHKVVEPGEQEHSFAFNIMEIQPFRRAAA
IQTMSLDAERRKLYVSSQWEVSQVPLDLCEVYGGGCHGCLMSRDPYCGWDQGRCISIY
SSERSVLQSINPAEPHKECPNPKPAPLQKVSLAPNSRYLSCPMESRHATYSWRHK
ENVEQSCEPGHQSPNCILFIENLTAQQYGHYFCEAQEGSYFREAQHWQLLPEDGIMAE
HLLGHACALAASLWLGVLPTLTLGLLVH"

BASE COUNT   1749 a   2358 c   2757 g   1997 t
ORIGIN

alignment_scores:
Quality: 1907.50   Length: 2757
Ratio: 3.222   Gaps: 18
Percent Similarity: 21.473   Percent Identity: 21.364

alignment_block:
US-09-041-236-2 x AF030697

Align seg 1/1: US-09-041-236-2 to AF030697 from: 1 to: 8861

    41 GluProHisThrValLeuPheHisGluProGlySerSerValTrpVa 57
      ||||::||||:||||||||||||||||||||||||||:||||||||:
    .1 GAGCCGCACACGGTGCTTTCCACGAGCCAGGCAGCCTCCTCGTGTGGT 50

    57 1GlyGlyArgGlyLysValTyrLeuPheAspPheProGluGlyLysAsnA 74
      ||||||||||||||||||||||||||||||||:||||||||||||||||
    51 GGGAGGACGTGGCAAGGTCACCTCTTTGACTTCCCCGAGGGCAAGAACG 100

    74 1aSerValArg............................... 77
      ||||||||||
   101 CATCTGCGCGCACGGTGAGCCTCTTCCTTCCCCAACACCCCCCCTACCC 150

  1293 GACAGGGGCACCATTCACAAGGTGGGGAATCAGGGGACCAGGACCATAG 1342

   416 rPheAlaPheAsnIleMetGluIleGlnProPheArgArgAlaAlaAlaI 433
      |||:||::||||:||||||||||::|||||||||:||:|||||||:|||
  1343 CTTTGCTTCAATATCATGGAGATCCAGCCCTTCACCGTGCAGCGGCCA 1392

   433 leGlnThrMetSerLeuAspAlaGluArgArgLysLeuTyrValSerSer 449
      ||||::|||||||||||||||||||||||||||||:::||:|
  1393 TCCAGCTATATCATTGGATGCTGACCGGCGGAAGCTCTATGTGACCTCC 1442

   450 GlnTrpGluValSerGlnValProLeuAspLeuCysGluValTyrGlyGl 466
      ||||||||||||||||||||||||||||||||||||||:::||
  1443 CAGTGGGAAGTGAGCCAGGTACCCCTGGACATGTGAGGGCTACAGCGG 1492

   466 yGlyCysHisGlyCysLeuMetSerArgAspProTyrCysGlyTrpAspG 483
      |||||||||||||||||||||||||||||||||||||||||||||||||
  1493 GGGCTGCCATGGCCGCTCATGTCCCGAGACCCCTACTGTGGCTGGGACC 1542

   483 lnGlyArgCysIleSerIleTyrSerSerGluArgSerValLeuGlnSer 499
      ||||::||:|||||||:::||||||||||||||||||||||
  1543 AGGACCGCCGCTGCGTGCTATCTACAGCTCCCAAGCAGTGCTGCAGTCC 1592

   500 IleAsnProAlaGluProHisLysGluCysProAsnProLysProAspLy 516
      |||:||:||||||||||||||::|||||||||:||||||||||||
  1593 ATTAATCCAGCGGAGCCACACAGAGTGTCCCAACCCTAAACCAGACGA 1642

   516 sAlaProLeuGlnLysValSerLeuAlaProAsnSerArgTyrLeuS 533
      :||||||||||||||||||||:::||||||||||||||||:::|||
  1643 GGCTCCATTGCAGAAGGTTTCCCTGGCCACTCTCGATACTACCTGA 1692

   533 erCysProMetGluSerArgHisAlaThrTyrSerTrpArgHisLysGlu 549
      ::||||||||||||||||||||||||||||||||||||||||||||
  1693 CCTGCCCATGGAGTCCCGCCACGCCACTTACTTATGGCGCCATGAGGAG 1742

   550 AsnValGluGlnSerCysGluProGlyHisGlnSerProAsnCysIleLe 566
      ||||||||||||||||||||||||||||:::||||||
  1743 AATGTGGAACAGAGCTGTGAGCCAGGCCACCAAAGCCCAGTCGCATCCT 1792

   566 uPheIleGluAsnLeuThrAlaGlnGlnTyrGlyHisTyrPheCysGluA 583
      ||||||||||||||||||||||||||||||:::||||||
  1793 GTTCATTGAGAACCTCACGGCCCGTGCAGTATGGCCACTACCGCTGCGAGG 1842

   583 laGlnGluGlySerTyrPheArgGluAlaGlnHisTrpGlnLeuLeuPro 599
      |||||||||||||||||||:::|||||||||||||||::||||
  1843 CCCAAGAGGGCTCCTACCTCCGTGAGGCTCAACACTGGGAGCTGCTGCCA 1892

   600 GluAspGlyIleMetAlaGluHisLeuLeuGlyHisAlaCysAlaLeuAl 616
      |||||||||::|||||||||||||:::|||| |||||||
  1893 GAGGACAGAGCACTGGCCGAACAACTAATGGGCCATGCCCGGGCCCTGGC 1942

   616 aAlaSerLeuTrpLeuGlyValLeuProThrLeuThrLeuGlyLeuLeuV 633
      ||||||||||||||||||||||||||||||||||||||||||||||||
  1943 CGCCTCCCTTCGGCTGGGGGTCCTGCCCACACTCATACTTGGTCTGCTGG 1992

   633 alHis 634
      |||||
  1993 TTCAC 1997

seq_name: gb_pr3:AF030697

seq_documentation_block:
LOCUS   AF030697   8861 bp   DNA   PRI   08-SEP-1998
DEFINITION Homo sapiens semaphorin L (SEMAL) gene, partial cds.
ACCESSION   AF030697
NID   g3551778
VERSION   AF030697.1   GI:3551778
KEYWORDS   .
SOURCE   human.
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 8861)
AUTHORS   Lange,C., Liehr,T., Goen,M., Gebhart,E., Fleckenstein,B. and

us-09-041-235-2.LEE.rge

Wed Jan 19 11:15:38 2000

```
123  .................................................  123
1051 CCCCATGTGCCTGATCAGCTCACCTTCTACTGCGTGGCCTTCTGCCCCTC  1100
123  .................................................  123
1101 ATGGTGGAAGGAGGATGGCGAGACTCCAATGCTGCCCTTGCCCTGGGAGG  1150
123  .................................................  123
1151 ATGGGGGCTCCTGGCCGAGAAACTGCCCGTCATGGGAGGCAGTGGCTGTGG  1200
124  ValAsnGlyThrV                                        128
     |||||||||||||
1201 GATTATGTGGCCATCCAACCCTCTGGATCTCCACAGGTGAATGGCACTG  1250
128  alValProLeuGlyMetArgGlyTyrAlaProPheSerProAspGlu     144
     ||||||||||||||||||||||||||||||||||||||||||||||
1251 TGGTGCCACTTGGCGGATGAGAGCTACGCCCCTTCAGCCCGGACGAG     1300
145  AsnSerLeuValLeuPheGlu.                               151
     ||||||||||||||||||
1301 AACTCCCTGGTTCTGTTTGAAGGTGGGCATGCTTCGGAACTGGGCTGG    1350
151  .................................................  151
1351 GAGCAGGATGGTCAGCTCTTTGTCCAGTGTCCGGAGGAGGACTTCCAGG   1400
152  GlyAspGlu                                           154
     |||||||
1401 AGCTGCCTGCCCTTACTCATTTCTCCCCCACTGACCCCGGGGACGAG     1450
155  ValTyrSerThrIleArgLysGlnGluTyrAsnGlyLysIleProArgPh  171
     |||||||||||||||||||||||||||||||||||||||||||||
1451 GTGTATTCCACCATCCGGAAGCAGGAATACAATGGGAAGATCCCTCGGTT  1500
171  eArgArgIleArgGlyGluSerGluLeuTyrThrSerAspThrValMetG  188
     .||||||||||||||||||||||||||||||||||||||||||||||
1501 CCGCCGCATCCGGGGCGAGAGTGAGCTGTACACCAGTGATACTGTCATGC  1550
188  lnAsn.                                              189
     ||:::
1551 AGAGTGAGTCAGGCTCCGGCTGGGCTGAGGGTGGCAAGGGGGTGTGAGC   1600
189  .................................................  189
1601 ACTTAAGGTGGCAGATGGGATCCTGATGTTCTGGAGGGCTCCCTGAGG    1650
189  .................................................  189
1651 GCCGCTGGGGCCATGCAGGAGAAAGCAGGAGACCTTGGTATAGGCTGGAGAAGT  1700
189  .................................................  189
1701 TAGGGTTGGCTGGGAGCAGCAGGAGGAACAGAACAAGGTATAGCAGTGGGATGGG  1750
189  .................................................  189
1751 CCCAGCCCTCTTCAGGGAACACAAAACAGAGGAGCCCCAGTGCAG       1800
189  .................................................  189
1801 GGTCCCCAGGAGCCAAAGTTATCCTCGTGAGTTCACGTGGAGGCAGC     1850
189  .................................................  189
1851 CCCCCAACTCCCTCTCCATCAGGGCTCTGCCAATTGAGCAGAAGTGACAT  1900
189  .................................................  189
1901 AGGGGCCCCAGGGACGACCTTCCCCCACTCCCCAGGCATGAAGTCATTGCTC  1950
189  .................................................  189
```

```
77   .................................................  77
151  TCTTATCTCCCCTCTGGCCTCCTCAGGGGTCCTCAGGAATCCGAGGGAG   200
77   .................................................  77
201  CTGGCTTCTTCCTAAACTGCCCCCACCTCCGTATCCTATAAATGGCTC    250
77   .................................................  77
251  CTGGGGAGGCTCCTAAAGGTAGTCCAGATTGGAGTGGGAGCTGGGGC     300
77   .................................................  77
301  GGTGTGGAGAAAAACAGGAGCTAATGGGCCTGGCCAGCTGGGCAGCCGCTG  350
77   .................................................  77
351  CTGCGGAAAGCCCAGGCTGGAAGCTGGGCCCCAGAGCCCATGCCTGGTCT   400
77   .................................................  77
401  TCTGAACCCTCTGGGCCTCAGCTCTGGATATGAGACCCGTGTTGACCTCA  450
77   .................................................  77
451  GGTAGATCACTCACCCCTTCAGAGCCCCCAGTTGCTCATCGTCTCAGATGAG 500
77   .................................................  77
501  AATAATGGTTGCTTCCTTTGGGGCTTATCCTGAGGCTGTGTGGGAAAGCAT 550
77   .................................................  77
551  TTCAGGGGTACCTCACCCCTGGCAGATTGAACTAATGCTTCTCCCCCTTCC  600
78   Thr.ValAsnIleGlySerThrLysGlySerCysLeuAspLysArg      92
     ||||||||||||||||||||||||
601  CCAGGTGAATATCGGCTCCACAAAGGGGTCCTGTCTGGATAAGCGGGTGA  650
92   .................................................  92
651  GCGGGGAGGGATCTGGAGGGGTCTGAGCCACTTGGTAAAGGGAGGAGGAG  700
92   .................................................  92
701  ACCCTGAGGGTCTAAGGAAGGAGCATGGCCCTGCCCCACGAGTCCCAGA   750
92   .................................................  92
751  CTGATGGGAGACCTGGTCCTCTCGTGCTTAGGGGATGGCGTCAGCTGAC   800
92   .................................................  92
801  ACACTCTGGGCTGTCCCGGAGGCTGTCACCTATGCTCTAAGCCCTTCTGAC  850
92   .................................................  92
851  ACCTTCTTCCCTGATCCTGGGGGTCCTAGTGCTTGCCAGGGGCCTT      900
93   AspCysG                                             95
     ||||||
901  CCAGCAACCAATTTCCTCCCTCTCCTTCTCCGGGCCAGGACTGCG       950
95   luAsnTyrIleThrLeuGluArgArgSerGluGlyLeuLeuAlaCys     111
     |||||||||||||||||||||||||||||||||||||||
951  AGAACTACACTCCTCGTGGAGAGGCCGGAGTGAGGGCTGCTGGCCCTGT   1000
112  GlyThrAsnAlaArgHisProSerCysTrpAsnLeu.               123
     ||||||||||||||||||||||||||||
1001 GGCACCAACGCCCGGCACCCCCAGCTGCTGGAACCTGGTGAGAAGGCTGCT 1050
```

```
1951 CTGGGCCGATGACATCTTTGTAGGAAGAGGGCAAAACAGGTGTGGGGTGG 2000
189  .................................................. 189

2001 AGGTGCAGGGTCTAGGGCCCCTCGGGAGTTGGACCTGATGTTATGAGTC 2050
189  .................................................. 189

2051 CTATTCCAGATCTGATTTGCCATGTGTTTGTGCAGACCCGAAGGAGGGAGG 2100
189  .................................................. 189

2101 AGAGTGTGCAGGGTTGGAATGGTCTCCCGGGCAAGCTTCCCAGCCTTACG 2150
190  ...........................Pro.GlnPheIleLysAlaThrIle 197

2151 CCCATTCGCTTCTGCGCTGCGACCCACAGTTCATCAAAGCCACCAT 2200
197  eValHisGlnAspGlnAlaAlaTyrAspAspLysIleTyrTyrPhePheArgG 214
     |||||||||||||||||||||||||||||||||||||||
2201 CGTGCACCAAGACCCTTACGATGACAAGATCTACTACTTCTTCCGAG 2250
214  luAspAsnProAspLysAsnProGluAlaProLeuAsnValSerArgVal 230
     ||||||||||||||||||||||||||||||||||||||||||
2251 AGGACAATCCTGACAAGAATCCTGAGGCTCCTCTCAATGTCCCGTGTG 2300
230  AlaGlnLeuCysArg. 235
     ||||||||||||||
2301 GCCCAGTTGTGCAGGTGAACACGGCGCGTGAGGGCTGCTGGCTACGTGTC 2350
235  .................................................. 235

2351 TGTGCATGAATAGGCCTGAGTGAGGGTGAGTTCTGTGTGTCCGTGTGTGCAT 2400
235  .................................................. 235

2401 GTAGAAGTTGTGTGGATGTATGAGTGGGTCTGTGTGTCAGGGACTGTGGGAG 2450
235  .................................................. 235

2451 CAGCTGTGTGTGTGCATGGAGCATCATGTGTCTGTGTGTGTGGGTAAAGGTGGC 2500
235  .................................................. 235

2501 TGAGCTCCTGTGCACGTATGATGCCGTGTGTGATGGCGTGTGTATGATGGGGTG 2550
235  .................................................. 235

2551 TGTGTGTGTGTGTGTGTGTTTGCCTGTGTGAATGTGCTGTGCCAC 2600
235  .................................................. 235

2601 GTATGTGGGTGCGTGAGTGCAGTGTCAGTAAATGTGTGTCTGAGTCCGTCTGCTCTG 2650
235  .................................................. 235

2651 TGGGGACCTGGCACTCACCTCACCTGCCCTGACCCTGCCACTGCTGCCCCTG 2700
235  .................................................. 235

2701 GGCTCTGGATCAGCCAGGCCTGCTTGCAGGAGTCTCATCTGGAGACCTGC 2750
235  .................................................. 235

2751 CCTGAGTCCTGGGGCACCCCGGCAGGTCCTGCCCCTCGGCCAGCCTGCCT 2800
235  .................................................. 235

2801 TCCTCCTCTCGGGCCCCAGGTGTTGATATTGCTGGCCAGTGGTTTCCTGGGGT 2850
235  .................................................. 235

2851 GTGTGGGAGAAGCCCGGCCAGTGTGTCAGGGGCCCTCTTCTCCCCCTCACCC 2900
236  ......GlyAspGlnGlyGlyGluserSerLeuSerValSerLysTrpAs 250

2901 TTCCAGGGGGACCAGGTGGGGGAAAGTTCACTGCAGTCTCCAAGTGGAA 2950
250  nThrPheLeuLysAlaMetLeuValCysSerAspAlaAlaThrAsnLysA 267
     ||||||||||||||||||||||||||||||||||||||||||
2951 CACTTTTCTGAAAGCCATGCTGGTATGCAGTGAGTGATGCTGCCAACAAGA 3000
267  snPheAsnArgLeuGlnAspValPheLeuLeuProAspProSerGlyGln 283
     ||||||||||||||||||||||||||||||||||||||||||
3001 ACTTCAACAGGCTGCAAGACGTCTTCCTGACCCCAGCCGGGCCAG 3050
284  TrpArgAspThrArgValTyrGlyValPheSerAsnProTrp....... 297
     ||||||||||||||||||||||||||||
3051 TGGAGGGACACCAGGGTCTATGGTGTTTTCCAACCCTGGTGAGTGGC 3100
297  .................................................. 297

3101 CCTTGTCCTGGGGCCCGGGCTGGCATTGGTCAGTGTCCAGTAGGGACAG 3150
297  .................................................. 297

3151 GAGGGCCTTGGGCCCTGCGAGGGCCTCCCTGGTGTGTGCAGGAGCAGGGGC 3200
297  .................................................. 297

3201 TGCAGGCTCAAGAGGCTGGGCTGTTGCTGGGTGTGGGGGACAGC 3250
297  .................................................. 297

3251 CAGTGCGATGTATGTACTGCTGTGTGAGTGAGTCTGCACTCATGGGTGTG 3300
297  .................................................. 297

3301 TGTGCATGCCCTATATGCACACTCATGACTGCCACTTGTGCCTGTGTGTCC 3350
297  .................................................. 297

3351 CACCACCTGCTTGTGCCGAGATGGACACTGGGCCCAGGAGGAAGCTGCT 3400
297  .................................................. 297

3401 GAAGCAATCTCTCGGGGAGCCTGGGTGCTGCTATTACACCTGCTCAGGCACTGCC 3450
297  .................................................. 297

3451 TGAGCCCGATAATTCACACTTCTTAATCACTCTCATTGATTGAACACACG 3500
297  .................................................. 297

3501 GCAGGCCGAAGTGTTGGGTGTGTGGGGAGAGTTAGGGATAGAGTGGAG 3550
297  .................................................. 297

3551 GAACCCAAGACCCTGCTCTGGGGTCCTGGGGTGAGTGGGTCCCCAGGCT 3600
297  .................................................. 297

3601 GGGAAGGGGTTGGGGTCTGCGCCTCCTGGGGCATCAGCACCCCACGCCT 3650
297  .................................................. 297

3651 GTGCCCAGGGAGGGCTTAGAGAACTGCTCAGCCTATGATGGGGTTCCTCCT 3700
297  .................................................. 297

3701 GCCTTGGGGTTGGGTAGAGCAGAGATGGCCTTCGAGACTTCAGTGATTCGTAA 3750
297  .................................................. 297

3751 CAGGATACAAGTTTGTGTTTTTAAATTGCAGCACAAAAGAAATTAGGCTGA 3800
```

```
 297    .................................................................    297
3801    ACTCCTCCTTCCTCTCCTCTCCATTTCAGTGGTGGTTGG                              3850
 297    .................................................................    297
3851    CAACTCAGTGCCAGGCACAAGGCTGGCCTGGGTGAGTGGAGGTGGATGGG                    3900
 298    .................................................   Asn.Ty           299
                                                         ::: ||
3901    TGGGTTCTGGGCCCCCATTGAGCTGGTCTCCATGTCACTGCAGGAACTA                     3950
 299    rSerAlaValCysValTyrSerLeuGlyAspIleAspLysValPheArgT                    316
        ||||||||||||||||||||||||||||||||||||||||||||||||||
3951    CTCAGCCGTCTGTGTATTCCCGTGACATTGACAAGGTCTTCCGTA                         4000
 316    hrSerSerLeuLysGlyTyrHisSerSerLeuProAsnProArgProGly                    332
        ||||||||||||||||||||||||||||||||||||||||||||||||||
4001    CCTCCTCACTCAAGGGCTACCACTCAAGCCTTCCAACCCGCGGCCCTGC                     4050
 332    .................................................                    332
4051    AAGGTGAGCGTGACACCAGCCGTGGCCCAGGCCCAGCCCTCCTTCTGCCT                    4100
 333    ..........................................Lys..CysLe                  335
                                                   ||||||
4101    CACCTCCCCACCCCACTGGGCTCTCGCTGCTGCTCGCTGCCAGTGCCT                      4150
 335    uProAspGlnGlnProIleProThrPheGlnThrValAlaAspArgH                      352
        |||||||||||||||||||||||||||||||||||||||||||||||
4151    CCCAGACCACAGCGATACCCCACAGAGACCTTCCAGGTGGCTGACCGTC                     4200
 352    isProGluAlaAlaGlnArgValGluProMetGlyMetGlyProLeuLysThrPro             368
        ||||||||||||||||||||||||||||||||||||||||||||||||||
4201    ACCCAGAGGTGGCGCAGAGGGTGGAGCCCATGGGCATGGGGCCTCTGAAGACGCCA              4250
 369    LeuPheHisSerLysTyrHisTyrGlnLysValAlaValAlaValHisArgMetGl             385
        ||||||||||||||||||||||||||||||||||||||||||||||||||
4251    TTGTTCCACTCTAAAATACCACTACCAGAAAGTGGCCGTCCACGCCATGCA                   4300
 385    nAlaSerHisGlyGluThrPheHisValLeuThrLeuThr                             398
        ||||||||||||||||||||||||||||||||
4301    AGCCAGCCACGGGGAGACCTTTCATGTGCTTTACCTAACTACAGGTGAGA                    4350
 398    ...............................................                      398
4351    GGCTACCCCGGGACCCTCAGTTTGCTTTGTGTAAAAACGGGCATGAAAGGTG                  4400
 398    ...............................................                      398
4401    TAAGGAATAAATGTAGTTAACATCTGGTTGGATCTTTACATGTGGAAGGAA                   4450
 398    ...............................................                      398
4451    TAATTGAGTGACTGGAGTTGTCAGGGGGTTAATGTGTGTGGTGTGGAAGA                    4500
 398    ...............................................                      398
4501    GCCAGGCAGGGAGCTTCCTGAGGAGGTAGGGGCAAGAGGGAAAAGGGG                      4550
 398    ...............................................                      398
4551    GATGGGGAGAAAAGCAAGCACTGGGATTTGGAGGCGGAAATCTGGAAGTC                    4600
 398    ...............................................                      398
4601    TGAGCAAAGCCAGGTGCACCTTTGGTCCAGATGTCTGACTCAGGGAAGAA                    4650
 398    ...............................................                      398
4651    GATGGTAGGAAGAGACCTGGCAAATGAGGAGGAGGGGCCTGAACCACAGG                    4700
 398    ...............................................                      398
4701    GATACTGGCCTCTGCCAGGCAGAATGAGGGAGTCAGGCGCCCTGCGCCTGTC                  4750
 398    ...............................................                      398
4751    TTTGGGATTGTGTGCAGGTGAGAAGAAACATTTGAGGAGTTGATGGGCACA                   4800
 398    ...............................................                      398
4801    AATTAGGTATGGGAAGGAGTTCCAGGGGCAGAACCTTTGCCATCTCAC                      4850
 398    ...............................................                      398
4851    AGAGGACAGGGGCAGCTTCTCTCTTCCCTGGAGTAGGCCCTGCTGGGGG                     4900
 398    ...............................................                      398
4901    AAGCTGGGTGGAATGCCGTGGGGAGATGCTCCTGCTTTCTGGAAAGCCACA                   4950
 398    ...............................................                      398
4951    GGACACGGAGGAGCCAGTCTGAGTTGGGTTTGTCGCAGCTTCCCATGCC                     5000
 398    ...............................................                      398
5001    AGCTGCCTTCCTTGGAGACTGGAAAGGGCCTCTAGCACCCCTGGGGCCATT                   5050
 398    ...............................................                      398
5051    CAATTCAGGCCCCAGGCGCCCAACCTCAGTTGTTCACATTCCCCATGTGAT                   5100
 398    ...............................................                      398
5101    CTCCTGTTGTCGTTCACCTTGGGACTGTCTCGGCTTTGGTGACCTTGTA                     5150
 398    ...............................................                      398
5151    GGAAACTGGAACCCCAGCACCATTGTTGGCTCCTGGAAGCCTTGGGGAG                     5200
 398    ...............................................                      398
5201    AGGAATTCCCACAGGGCAGGGCCTGGGCTCTGATTCCCTGCCGCCTCTTTAC                  5250
 398    ...............................................                      398
5251    TCCCTATTCATCCCGGCTACACCCTTGGGCCCCATCCTTGCTTGGCTCC                     5300
 398    ...............................................                      398
5301    AGTACTGGCTGGCCACAGCTGTTGTGGCTCATCCAGGGATGGCAGGGCACTG                  5350
 398    ...............................................                      398
5351    GGGAACGAAGAGAGAGGTCACACAGTGCGGAACTGGAGGAGGAGCTAG                      5400
 398    ...............................................                      398
5401    GACAAGGAAGGCTGGACTTGGGCCATGGGCCATGGATGATTCCCTTCCTGCAGACTTGGG         5450
 398    ...............................................                      398
5451    AAGTGAGCACACTTGAGTGATTAGAGAAGGTGTCTTCGTTCTAAGGGCAG                    5500
 398    ...............................................                      398
5501    TGGAGGAGGCACCATTTGGAGCCTGCATCATCGTATTTGGGCTAGATT                     5550
 398    ...............................................                      398
5551    GAAAAATAGAGCTTCTAAGTCCTTCTGCAGAGAATGGGGAGGCTCTCACAA                   5600
 398    ...............................................                      398
```

us-09-041-236-2.LEE.rge

```
5601 CTGGGAGAAGTATTGGCTCTTTCCTGAGAGAATTTGCCAAGGGTATGCTG 5650
 398 ..................................................
5651 TTACTGGGGCTGGTTGGGAAGGAGTATAGGGCATTATGTGTGTGAAGGCA 5700
 398 ..................................................
5701 GTGGCTGGGTGGGGGGCCTTATCAGGCCCCAAGGAGCATCTGGCCACATCTC 5750
 398 ..................................................
5751 AGAGTCCACAGATGAGGATCACGGATGTGTAGAGGAAACATCCTAGGCAG 5800
 398 ..................................................
5801 GCAATCATCTGACTGCTTTTTGGGGCAGTGCCCTGGGAAATTGGG 5850
 398 ..................................................
5851 AGGGAGGGAGAGGAGGGAGGTAGGCTATTCTAGAAACTGGGAGAGCAGGTG 5900
 398 ..................................................
5901 AGTTAGGATTGGGAGGACCAGGGGTCAGGGGTCCCCATTGGTCCCTAATTG 5950
 398 ..................................................
5951 AGAACGGAGAGAGCATTGGTCTAGGAGGCAGGCAGCTCGGTCGGTATAAGACC 6000
 398 ..................................................
6001 TTGGGAACTCTTGATTTAGAATCCAAGATCCTTTTTAGATCTAGGATTT 6050
 398 ..................................................
6051 ATAAAATTAAGATATCCCCTAAGATCAAATGCAACGTGGAGTCCTGAATT 6100
 398 ..................................................
6101 GGATCCTAGAACACAGAAGAGGACATTTGTTGTGGAAAAACTAGTGAAATCCAA 6150
 398 ..................................................
6151 ATAAAGTCTGTAGTTTGTTAATAGTAATGCACCAAATGTCAGTTGCCTAG 6200
 398 ..................................................
6201 TTGTGACAAATATACCGTGGTATGTAATGTAAGATGTAACATTAGGGGGAACT 6250
 398 ..................................................
6251 GGAGAAGGGTAGATTGGAGCTCTCTGTACTATCTTGCAACTTTTCTGGG 6300
 398 ..................................................
6301 AATCTAAAATTACTCCAAAATAAAAAAAAAAAATGTATTTAAAGTAAATATA 6350
 398 ..................................................
6351 TTCCCTAAGAGTCCAGGAGGCAGGGGAGTGTAGAGCAGCTGAGTGGTT 6400
 398 ..................................................
6401 GGGTTCTGACAGATTTGGTTCCAACTCGGTCTCTGCTCACCAGCTGT 6450
 398 ..................................................
6451 GTGACCTTGAGCAAGTGGCTTAGCCTTTCTGAGCCTGATTTCCTTATCTG 6500
 398 ..................................................
```

```
6501 TGGAGTGGGAAGATGACAGCCACCTCGCGCAGGGCGTGTGGAGGGTTAAACG 6550
 398 ..................................................
6551 AGGTGATGCATGGCAGCAGCCGCACTGACCTTGCTGGTGTGGGGCTCCT 6600
 398 ..................................................
6601 GCTTCTTGTTCTTCCCGTGCAGCCTTGGGAATGTTGAGGCCGTATCCAGG 6650
 398 ..................................................
6651 GACCCCTGGCGCTCCTGGGATGGCCTCTCTGGATCAGCCTTGGAAGGTTC 6700
 398 ..................................................
6701 CAGGCTGCCCTTAGGCTCCCACATTCTCCCAGTCACGCTCTCCTCGCC 6750
 398 ..................................................
6751 CTGCCCACACCAGTCCTGTGACCCTTGCCTGAGTTGTGACTTCCCACCCC 6800
 398 ..................................................
6801 TCCCCGGCCTAGAGGAAAGCTGCCTGGCCCGTCCCAGTGGACTCCCGCCCA 6850
 398 ..................................................
                 Thr. AspArgGlyThrIleHisLysValVa 408
6851 CTGACCCTCTGTCCACCATACACAGACAGGGGCACTATCCACAAGGTGGT 6900
        ||  ||||||||||||||||||||||||||||||
 399                                          Arg.
                 | |
        lGluProGlyGluGlnGluHisSerPheAlaPheAsnIleMetGluIleG 425
        ||||||||||||||||||||||||||||||||||||||||||||||
6901 GGAACCGGGGAGCAGGAGCACAGCTTCGCCTTCAACATCATGGAGATCC 6950
 408
        lnProPheArgArgAlaAlaAlaIleGlnThrMetSerLeuAspAlaGlu 441
        ||||||||||||||||||||||||||||||||||||||||||||||||
6951 AGCCCTTCCGCCGCGCGGCGGCCATCCAGACCATGTCGCTGGATGCTGAG 7000
 425
        Arg.                                           442
        | |
7001 CGGGTGCAGCCTTCCCCACTCGCGTCCCATGGGCTATGCAGTGACTGCAGC 7050
 442
7051 TGAGGACAGGGCTCCTTTGCATGTGATTGTGTGTCTTTTAAGAGCTTC 7100
 442
                                   ArgLysLeuTyrValSerSerGl 450
7101 TAGGGCCTTAGGGCCTGGACATTAGGAGTCGAGTGTGGGGTGGGCCCGGG 7150
 443
        CCTGACCAATCCTGCTGTCCTTCCAGGAGGACGTGTATGTGAGCTCCCA 7200
 450 nTrpGluValSerGlnValProLeuAspCysGluValTyrGlyGlyG 467
7201 GTGGGAGGTGAGCCAGGTGCCCCTGGACCTGTGTGAGGTCTATGGCGGGG 7250
        lyCysHisGlyCysLeuMetSerArgArgProTyrCysGlyTrpAspGln 483
        ||||||||||||||||||||||||||||||||||||||||||||||||
7251 GCTGCCACGGTTGCCTCATGTCCCGAGACCCCTACTGCGGCTGGGACCAG 7300
 467
        GlyArgCysIleSerIleTyrSerGluArg 494
        ||||||||||||||||||||||||||
7301 GGCCGCTGCATCTCCATCTACAGCGAACGTAGTTGCCGGGATCC 7350
 484
7351 CTCCGTCCCTGGGACAAGGTGGGCATGGGACAGGGGGAGGTGTTGTCGGG 7400
 494
7401 CTGGAGAGGTGGCCGACTGGGCCTTTCTTGTGGACCTCCTCTCTACT 7450
 494
```

```
DEFINITION   Mus musculus semaphorin L (SemaL) mRNA, partial cds.
ACCESSION    AF030699
NID          g3523116
VERSION      AF030699.1  GI:3523116
KEYWORDS     .
SOURCE       house mouse.
  ORGANISM   Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
             Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE    1  (bases 1 to 1192)
  AUTHORS    Lange,C., Liehr,T., Goen,M., Gebhart,E., Fleckenstein,B. and
             Ensser,A.
  TITLE      New eukaryotic semaphorins with close homology to semaphorins of
             DNA viruses
  JOURNAL    Genomics 51 (3), 340-350 (1998)
  MEDLINE    98389619
REFERENCE    2  (bases 1 to 1192)
  AUTHORS    Ensser,A., Liehr,T., Lange,C., Goen,M., Gebhart,E. and
             Fleckenstein,B.
  TITLE      Direct Submission
  JOURNAL    Submitted (21-OCT-1997) Institut fuer Klinische und Molekulare
             Virologie, Friedrich Alexander Universitaet, Schlossgarten 4,
             Erlangen, Bavaria 91054, Germany
FEATURES             Location/Qualifiers
     source          1..1192
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /map="9A3.3-B"
                     /chromosome="9"
     gene            1..>1192
                     /gene="SemaL"
                     /evidence=not_experimental
     CDS             13..>1192
                     /gene="SemaL"
                     /note="M-Sema-L"
                     /codon_start=1
                     /evidence=not_experimental
                     /product="semaphorin L"
                     /protein_id="AAC34262.1"
                     /db_xref="PID:g3523117"
                     /db_xref="GI:3523117"
                     /translation="MTPPPGRAAPSAPRARVLSLPARFGLPLRLRLLLVFWVAAASA
                     QGHSRSGPRISAVWKGQDHVDFSQPEHTVLFHEPGSFSVWVGGRGKVYHFNFPEGKN
                     ASVRTVNIGSTKGSCQDKQDCGNYITLLERRGNGLLVCGTNARKPSCWNLVNDSVVMS
                     LGEMKGYAPFSPDENSLVLFEGDEVYSTIRKQEYNGKIPRFRRIRGESELYTSDTVMQ
                     NPQFIKATIVHQDQAYDDKIYYFFREDNPDRKNPEAPLNVSRVAQLCRGDQGGESSLSV
                     SKWNTFLKAMLVCSDAATNRNFNRLQDVFLLPDPSGQWRDTRVYGVFSNPWNYSAVCV
                     YSLGDIDRVFRTSSLKGYHMGLSNPRPGMCLPKKQPIPTETFQVADSHPEVAQRVEPM
                     G"

BASE COUNT      254 a      355 c      336 g      247 t
ORIGIN
```

```
alignment_scores:
               Quality:    1748.00             Length:       364
               Ratio:      5.067               Gaps:         2
Percent Similarity:  94.780     Percent Identity:  90.385

alignment_block:
US-09-041-236-2 x AF030699

Align seg 1/1    to: AF030699       from: 1   to: 1192

    1 LeuLeuLeuLeuTrpAlaAlaAlaAlaSerAlaGlnGlyHisLeuAr   17
      ||||||||||||||| ||  ||   |||  |||||  ||||||||| ||
  109 CTTCTGCTGGTGTTCTGGGTGGCCGCCGCCTCCGCCCAAGGCCACTCGAG  158

   17 gSerGlyProArgIlePheAlaValTrpLysGlyHisValGlyGlnAspA   34
      ||||||||||||||| |||||||||||  |||||||||||||||||:::||
  159 GAGCGGAGCCCCGACATCTCCGCCGTCTGGAAA........GGGCAGGACC  199

   34 rgValAspPheGlyThrGlnThrGlyProHisThrValLeuPheHisGluPro   50
      ::||||||||||||:::|||   |||||||||||||||||||||||||
```

```
  494 ...................................................  494
 7451 GGAACTGCACTAGGGGTAAGGATAATGAGGGTCAGGTCTGCAGCCTTGTAT  7500

  495 ...........................Ser..ValLeuGlnSerIl  500
                                 |||  :::||||||
 7501 CTGCTGATCCTCTTCGTCCCTTCCCACTCCAGTCAGGTCAGTGCTGCAATCCAT  7550

  500 eAsnProAlaGluProHisLysGluCysProAsnProLys.........  513
      ||||||||||||||||||||||||||||||||||
 7551 TAATCCAGCCGAGCCACAACAAGGAGTGTCCCAAACCCAGGTACCT  7600

  513 ..................................................  513

 7601 GATCTGGCCCTGCTGGGCGTGTGGCCCAATGAGTGGGGTACTGCCCTGC  7650

  513 ..................................................  513

 7651 CCTGATTGTCCTGGTCTGAGGGAAACATGGCCTTGTCCTGTGGCCCCAG  7700

  513 ..................................................  513

 7701 GTACATGGGGCAGGATACAGTCCTGCAGAGGGAGCCCTCTTGGTGGGATG  7750

  513 ..................................................  513

 7751 AGCGAGACGGGAGAAAAAAGGAGGACGCTGAGGGTGTTGGGTTCCCACGTT  7800

  513 ..................................................  513

 7801 CATTCAGAAGCCCTTGTCCTGGATCCCAGTCCGTCGGGGAGGACACATCCT  7850

  514 ...............................................P  514
                                               |
 7851 CCCCTGGGAGCTCTTTGTCCCTCCTCACGGCTGCTTCCCCCACTGCCTCCC  7900

  514 rOAspLysAlaProLeuGlnLysValSerLeuAlaProAsnSerArgTyr  530
      |||||||||||||||||||||||||||||||||||||||||||||||||
 7901 CAGACAAGGCCCCACTGCAGAAGGTTTCCCTGCCCCAAACTCTCGCTAC  7950

  531 TyrLeuSerCysProMetGluSerArgHisAlaThrTyrSerTrpArgHi  547
      |||||||||||||||||||||||||||||||||||||||||||||||||
 7951 TACCTGAGCTGCCCCATGGAATCCCGGCACGCCACCTACTCATGGCGCCA  8000

  547 sLysGlyAsnValGluGlnSerCysGluProGlyHisGlnSerProAsnC  564
      |||||||||||||||||||||||||||||||||||||||||||||||||
 8001 CAAGGGAAACGTGGAGCAGAGCTGCGAACCTGGTCACCAGAGCCCCAACT  8050

  564 ysIleLeuPheIleGluAsnLeuThrAlaGlnGlnTyrGlyHisTyrPhe  580
      |||||||||||||||||||||||||||||||||||||||||||||||||
 8051 GCATCCTGTTCATCGAGAACCTCACGGCCCAGCAGTACGGCCACTACTTC  8100

  581 CysGluAlaGlnGluGlySerTyrPheArgGluAlaGlnHisTrpGlnLe  597
      |||||||||||||||||||||||||||||||||||||||||||||||||
 8101 TGCGAGGCCCAGGAGGGCTCCTACTTCCGCGAGGCTCAGCACTGGCAGCT  8150

  597 uLeuProGluAspValGlyIleMetAlaGluHisLeuLeuGlyHisAlaCysA  614
      |||||||||||||||||||||||||||||||||||||||||||||||||
 8151 GCTGCCCGAGGACGGCATCATGGCCGAGCACCTGCTGGGTCATGCCTGTG  8200

  614 laLeuAlaAlaSerLeuTrpLeuGlyValValProThrLeuThrLeuGly  630
      |||||||||||||||||||||||||||||||||||||||||||||||||
 8201 CCCTGGCCGCCTCCCTGTGGCTGGGGGTGCTGCCCACCCTCACTCTTGGC  8250

  631 LeuLeuValHis  634
      ||||||
 8251 TTGCTGGTCCAC  8262 ,
```

```
seq_name: gb_ro:AF030699

seq_documentation_block:
LOCUS       AF030699       1192 bp    mRNA              ROD       03-SEP-1998
```

```
                                           350 AspArgHisProGluValAlaGlnArgValGluProMetGly  363
                                               |||  |||||||||||||||||||||||||||||||||||||||||
                                          1150 GATAGTCACCCAGAGGTGGCTCAGAGGGTGGAACCCATGGGG  1191

seq_name: gb_vi:AF005370

seq_documentation_block:
LOCUS        AF005370   130608 bp   DNA           VRL      21-AUG-1997
DEFINITION   Alcelaphine herpesvirus 1 L-DNA, complete sequence.
ACCESSION    AF005370
NID          g2337967
VERSION      AF005370.1   GI:2337967
KEYWORDS     .
SOURCE       wildebeest.herpesvirus.
  ORGANISM   Alcelaphine herpesvirus 1
             Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
             Gammaherpesvirinae.
REFERENCE    1  (bases 1 to 130608)
  AUTHORS    Ensser,A., Pflanz,R. and Fleckenstein,B.
  TITLE      Primary structure of the alcelaphine herpesvirus 1 genome
  JOURNAL    J. Virol. 71 (9), 6517-6525 (1997)
  MEDLINE    97404659
REFERENCE    2  (bases 1 to 130608)
  AUTHORS    Ensser,A., Pflanz,R. and Fleckenstein,B.
  TITLE      Direct Submission
  JOURNAL    Submitted (27-MAY-1997) Institut fuer Klinische und Molekulare
             Virologie, Friedrich-Alexander-Universitaet Erlangen-Nuernberg,
             schlossgarten 4, Erlangen 91054, Germany
FEATURES             Location/Qualifiers
     source          1..130608
                     /organism="Alcelaphine herpesvirus 1"
                     /strain="C500"
                     /specific_host="Connochaetes taurinus taurinus"
                     /db_xref="taxon:35252"
                     /lab_host="MDBK bovine epithelial kidney cells; ATCC
                     CCL-22"
     CDS             918..1211
                     /note="putative ORF"
                     /codon_start=1
                     /evidence=not_experimental
                     /product="A1"
                     /protein_id="AAC58052.1"
                     /db_xref="PID:g2337968"
                     /db_xref="GI:2337968"
                     /translation="MRLHEFLFYRSIEFMYKKSRHKIFCITLSLYFLTAHKRPSARNP
                     ITDFFSRIWGPCFYWSRHQLLVNQLMLSPCRNKCSLTSDQCITANEDTK"
     polyA_signal    1248..1254
                     /evidence=not_experimental
     polyA_signal    complement(1300..1305)
                     /evidence=not_experimental
     CDS             complement(join(1796..2230,2313..2477))
                     /note="spliced cDNA detected by 5'RACE-PCR"
                     /codon_start=1
                     /evidence=experimental
                     /product="A2"
                     /protein_id="AAC58053.1"
                     /db_xref="PID:g2337969"
                     /db_xref="GI:2337969"
                     /translation="MSQNSNSENPSPRKKRYVKMCDLTEEQKERRSINRRASKNFLK
                     RRRIPEEQDEKGLINLKYENSRLRCQVERRKDEIRILREWLNYHKCTTLQNYNTGPPE
                     PRVKVENSLEMQCATAFLNLDQQYTTNNLNIPETVSGNNTTNGFAAATATLHTNCYEK
                     TLANNTNNFEAKLNCEVLPSFTSALDDLLSIDWNNLYNL"
     CDS             3492..5453
                     /note="A3; AHV-sema, similar to Vaccinia A39"
                     /codon_start=1
                     /evidence=not_experimental
                     /product="semaphorin homolog"
                     /protein_id="AAC58054.1"
                     /db_xref="PID:g2337970"
                     /db_xref="GI:2337970"
                     /translation="MAYLNATVSKPVISLLSLSKVLKFEHCGGEGQCLGLITEFVIH
                     PAAMGTLCVSIRLLMILSAITAAKSRFIDKPRLIYNLTDGFGQHRFFGPQEPHTVLFH
                     SLNSSDVYVGGNNTYLFDFAHSSNASTALINITSTHNTHRLSSTCENFITLLHNQTD
```

```
 200 ATGTGGACTTTAGCCAGCCTGAGCCACACCGTGCTTTCCATGAGCCG  249
                                                    |||||||||||||||||||||||||||||||||||||||||||||||
  51 GlySerSerValTrpValGlyLysAsnAlaSerValTyrLeuPheAs  67
     ::|||||||||:::||::||||||:||||:||||||||    |||::
 250 GGCAGCTTCTCTGTCTGGGTGGGTGGCAAGGTCTACCACTTCAA  299

  67 pPheProGluGlyLysAsnAlaSerValArgThrValAsnIleGlySerT  84
     :||||||||||||||||||||||||||||||||||||||||||||||||
 300 CTTCCCGGAGGGCAAGAATGCCTCTGTGCCACGGTGAACATCGGCTCCA  349

  84 hrLysGlyserCysLeuAspLysArgAspCysGluAsnTyrIleThrLeu  100
     ||||||||||||  |||||||||||||||||||||||| |||::|||||
 350 CAAAGGGCTCCTGCAGGACAAACAGGACTGTGGGAATTACATCACTCTT  399

 101 LeuGluArgSerGluGlyLeuLeuAlaCysGlyThrAsnAlaArgHi  117
     ||||||||||||||||||||||||:|||||||||||||||||||||||
 400 CTAGAAAGGGCGGGTAATGGCGTGCTGGTCTGTGGCACCAATGCCCGGA  449

 117 sProSerCysTrpAspAsnLeuValAsnGlyThrValVal...ProLeuGlyG  133
     :|||||||||||||||||||||||||    ::|||||||||||||||
 450 GCCCAGCTGCTGGAACTTGGTGAATGACAGTGTGGTGATGTCACTTGGTG  499

 133 luMetThrArgGlyTyrAlaProPheSerProAspGluAsnSerLeuValLeu  149
     |||||::||||||||||||||||||||||||||||||||||||||||
 500 AGATGAAAGGCTATGCCCCTTCAGCCCGGATGAGAACTCCCTGGTTCTG  549

 150 PheGluGlyAspValTyrSerThrIleArgLysGlnGluTyrAsnGl  166
     |||||||||||||||||||||||||||||||||||||||||||||||
 550 TTTGAAGGAGATGAAGTGTACTCTACCATCCGGAAGCAGGAATACAACGG  599

 166 yLysIleProArgPheArgArgIleArgGlyGlySerGluLeuTyrThrS  183
     ||||||||||||||||||||||||||||||||||||||||||||||||
 600 GAAGATCCCTCGGTTTCGACGAATTCGGGGCGGGAGTGAACTGTACACAA  649

 183 erAspThrValMetGlnAsnProGlnPheIleLysSerAlaThrIleValHis  199
     |||||||||||||||||||||||||||||||||||||||||||||||||
 650 GTGATACAGTCATGCAGAACCCCACAGTTCATCAAGGCCACCATTGTGCAC  699

 200 GlnAspGlnAlaTyrAspAspLysIleTyrTyrPhePheArgGluAspAspAsPAs  216
     |||||||||||||||||||||||||||||||||||||||||||||||||
 700 CAAGACCAAGCTATGATGATGATAAGATCTACTACTTCTTCCGAGAAGACAA  749

 216 nProAspLysAsnProGluAlaProLeuAsnValSerArgValAlaAlaGlnL  233
     |||||||||||||||||||||||||||||||||||||||::|||||||
 750 CCCTGACAAGAACCCCGAGGCTCCTCTCAATGTGTCCAGAGTAGCCCAGT  799

 233 euCysArgGlyAspGlnGlyGlyGluSerLeuSerValSerLysTrp  249
     |||||||||||||||||||||||||||||||||||||||||||||||
 800 TGTGCAGGGGAGACCAGGGTGGTGAGAGTTCGTTGCTGTCTCCAAGTGG  849

 250 AsnThrPheLeuLysAlaMetLeuValCysSerAspAlaAlaThrAsnLy  266
     |||||||||||||||||||||::||||||||||||||||||||:::||
 850 AACACCTTCCTGAAAGCCATGTTGGTCTGCAGCGATGCAGCCACCAACAG  899

 266 sAsnPheAsnArgLeuLeuAsnProValPheLeuProAspProSerGlyG  283
     :|||||||||||||||||||||||||||||||||||||||||||||||
 900 GAACTTCAATCGGCTGCTAAATCCGGTATTGCTTCCTGACCCCAGTGGCC  949

 283 lnTrpArgAspThrArgValThrGlyIleValPheSerArgAsnProTrpAsnTyr  299
     |::||||||||||||||||||||||||||||||||||||||||||||||
 950 AGTGGAGAGATACCAGGGTCACCGGGTATGGCGTTTTCTCGAAACCCTGGAACTAC  999

 300 SerAlaValCysValTyrSerLeuGlyAspIleAspSerValPheArgTh  316
     :|||||||||||||||||||||||||||||::||||||||||||||||
1000 TCAGCTGTCTGCGTGTATTCGTTGGTTGCTACATTGACAGAGTCTTCCGTAC  1049

 316 rSerSerLeuLysGlyTyrHisSerSerLeuProAsnProArgProGlyLeuL  333
     ||||||||||||||    :::||||||||||||||||||||||||||||
1050 CTCATCGCTCCAAAGGCTACCACCATGGCCCTTTCCAACCCCTCGACCTGGCA  1099

 333 ysCysLeuProAspProGlnProIleProThrGluThrPheGlnValValAla  349
     ||||||||  :::|||||||||||||||||||||||||||||||||||||
1100 TGTGCCTCCCAAAAAAGCAGCCCATACCCACCAGAAACCTTCCAGTAGCT  1149
```

```
GLLACGTNSQKPSCWLINNLTTQFLGPKLGLAPFSPSSGNLVLFDQNDTYSTINLYKS
LSGSHKFRRIAGQVELYTSDTAMHRPQFVQATAVRHNESYDDKIYFFQENSHSDFKQ
FPHTVPRVGQVCSSDQGGESSLSVYKWTTFLKARLACVDYDTGRIYNELQDIFIWQAP
ENSWEETLIYGLFLSPWNFSAVCVFTVKDIDHVFKTSKLKNYHHKLPTPRPGQCMKNH
QHVPTETFQFVEVADPVYQKNNAMFPIQSKYIYTKLLVYRVEYGGVFWATIFY
LTTIKGTIHIYVRYEDSNSTTALNILEINPFQKPAPIQNILLDNPNLKLYVNSEWEVS
EVPLDCCSVYGNDCFSCFMSRDPLCTWYNNTCSFKQRVSVETGGPANRTLSEMCGDHY
APTVVNRDNYILAGSYLSCPAVSNHADYFWTKDGFTEKRCHVKTHKNDCILLIANS
TTATNGTHVCNMKEDSVTKLLEVNVTLM"

polyA_signal    5590..5597
CDS             5732..6097
                /note="ORF containing putative signal peptide; detected by
                5'RACE-PCR"
                /codon_start=1
                /evidence=experimental
                /product="A4"
                /protein_id="AAC58055.1"
                /db_xref="PID:g2337971"
                /db_xref="GI:2337971"
                /translation="MVAQLYHHLTLIASIYIIFFVNAAPTLYEDDEEDFNSTEYDVF
                LDSDSFQFPRRNSSCNCTREEATRTLLEALKRALQLIIAGYINETDTEELPTYPPTMT
                PLETTPLDTSPPVLFSAIP"

polyA_signal    6123..6128
CDS             6333..10442
                /note="ORF03; similar to H. salmiri and EHV2 ORF3, similar
                to ORF75"
                /codon_start=1
                /evidence=not_experimental
                /product="tegument protein"
                /protein_id="AAC58056.1"
                /db_xref="PID:g2337972"
                /db_xref="GI:2337972"
                /translation="MITRELQLLAYTAEPSALETAAIAALRSIPGLQNVIIQTQDAYL
                VTFFTSPRPLREHQLKIETLLFIKAALRTFDEQQYLPIPLSALSSSFTFVYGPDINRL
                PTSQSNELTAILQTRDAREKVFNVERIQHCRLLFSGPGSEQLNHTHYALLREILCGD
                LTLYQPFRNELFNSSLDFYATIYPPRFLKDRVRHVHGDINLNIIGAASNYNWASPHT
                AVTDSEIFRIPYTHVVRFIDRWEIRVRVPDVGFHGSRGVRRFQHGGAAYMCSTLNSTT
                GATGCLQKGMITSFISQPTLGOLGLITEPMCGWGQOELSATAIQMLTKYANTVEAFAK
                TLNYAGVPLVQGFVTLSPTLREKRAMAFTGSLSITGLPPYILKPPTAETLDAAMERNR
                QLLLVEVGYPDYGKGKLHNPVNIMNSEGRHAHILHQALRQLMIITPFGKVIHICCDW
                QYNEKVTMFKIAKACGDLGLSMSARSLPAHTLRLKKWNFPRNLNYNHRIIKNSWLKVD
                SAAALVVIADNEYQDVSKKMDIALSGWGCPFHILGNLTPNSNTIVISDKNQYGEIVDI
                QYKMHMPKQPSEGTEDTPLAPDMSNIQLFKNLDVTEDLLLQVLRHPTVGCKAHIVHHV
                DRCGNGHIAQQPGVCPFFDIPLCDFSVTVHNLVDGDIREGMESVPRVWAADWRVARRLI
                ETQYSTPGLIDTATLANNLGLTYHIPSEHQVTVESKKYGNCIGIGEKTFTQRDPLL
                GTILAIVESCTNCILGPVNMEIFLIGLSIVPEGIHYRHEVNSIMAMAKDFCSSMNF
                GFQVNSAENGNCLLRSVVATANAPCVVPGPSLKPYFKKPGSAILRVNLHTEHFLSGGI
                CCMASGIGASETFTPTPSQSLRNLLQFMLIVKAENLALSGHDVSDGLICAVCEMMFAG
                GLSARLIHDDEEPVFLFSETVNQVLVEAFVDYAIFRGIIARANLINVKAKGSFTLAI
                DTFTVFHQNTQLLSVPVSRLKHNWTLFSKSVDLLVVKEDQVLPEETSYGNYEVHLTVD
                PYSIISQSTRRPNVLHLLPGCGYPDALLAALTNSGFSPDIVVYPGCKYRHNRREDAA
                PGSPTADDFIAGIVLYGSSNIDSDVGDSTIRQWLNVNRQVINDVRRNLKAKGSFTLAI
                GQLACRILFATKAIGFDAGSQQTPFLLPNASRRYESRWLNFKIPEDTKAVAFRDLRGC
                VLPCWVQGTHLGFSHNNITFGDLETRQQVAATFNGPLVQSGPAREYPLNPTEAEPHY
                AGLCSDGRHLALLFDPCLAFNTWQWHQNQTGPGQGELPVSPWKLMFYRLYNWSKFHQ
                HYRSLLRTNLRHTFNFETQQLDFPHADRRGAVPHDNPQYIPMDAMDPSQ"

polyA_signal    11134..11139
CDS             11260..14643
                /note="ORF06"
                /codon_start=1
                /evidence=not_experimental
                /product="major ss DNA binding protein"
                /protein_id="AAC58057.1"
                /db_xref="PID:g2337973"
                /db_xref="GI:2337973"
                /translation="MALKHQLNHAVEDNLGSKAPIGPCGFIYIYPETHFNFKEISLLG
                DKYAEAGAFSLPLLHGVTVEEAFVPNVKAVYKKIDMTTVSVKLSTFYNRAIIFHNVEK
                FESIFSGPGLGSLCKEACDLFGYVPFTLGEGSTDVSDICPPVWQEKDAIMAVVITEG
                FKERLHLGKLYLKSOMHSVMINKTEVYRIPLYDEDLFTKKSSLRRLYLPAVSEYLY
                TLYTSLAQSLRVHNAASLVEAIQEQFVHDKYKMAKLVSFKEYPLATVGACDTTLMVID
                AVAAELGLSYSLSFFEAPQEKTKVQDYYSWDIFASCETDSDRLEALSKWNALQAIHIH"
```

```
AQLFSTNSIYYNRVARQAPIPNSKVEPNVINSYLQHGLANLCEETLFEDGSPAFTG
APASSLDGSSFTLQHLAYAAAFSPNLLARMCIYLQFCQHQKSTLNPAYNITEYVGSAA
NSPVCSLCSGQCPCVCINTLFYRLKDRFPPVLQGSRRDPYVITGINVENELDFLGNF
ASFRDKDEDQNQTEETPRYTWQLNQTLTEKLEAAGLVDSPVADEGAGGSGSMNLEKF
VRTFSDIDSLVDAEAAKFINTMIKNNVMFKESIGVSHVIQYNCNTWQAPCSLMLNL
YYRSILTIIQDIALPISTVYESENPAQGYKPNEWLKLHYQLMTNFKSFFIDKGVITG
TEMKVVHAEQFSDFFDVDAAITNMYSFVPKVQVRLARAQVLALKNIKVKNRILFSGTSM
SEHYQNAFLKTANRRDNYILAGPVYKFTDGFPNLKISCLYLMSNFCKKKQIPC
VPGVSAEALNKFFSYINNSKQFEEVNMLDVVPDSYVTTYAKQRLNNAILRACGQTQFY
AVTIHSIFPKVQETCALEYPHVLGTSSVDSVEDYVNNVQNLKALTVNSSLRESAANLA
RSRPIVTLPVVVNKYTGIAGNALQFQSANLGYFMGRGVDKNLLGDSLFVKKQQNSYMR
KKYLFMTPLVGNLLKPSYTHQCTAFEIETVKRTIQSILEDQADEDVLNRVVCELVKSL
GAGCADLTLDDIQFYLGSYGMFSENILEKLDQLRELVGPWTHEWAESVLKSGTCETDE
VQFVAFEEQVKLTSMDHSGKVVGGKKRKIATMFDDLDL"
                14725..16767
                /note="ORF07"
                /codon_start=1
                /evidence=not_experimental
                /product="transport protein"
                /protein_id="AAC58058.1"
                /db_xref="PID:g2337974"
                /db_xref="GI:2337974"
                /translation="MGQYLAALYSQIYGLCLDVSLVEFCKPTSLCLTKIADACNKVHK
                IHEYVASLLQQNSLEACALSELSHLLENLKTRLFFIYHALLDNPTYFSKLHSSVGL
                CDLHKKLNVQFYNECGIEVNLTLINDIERFLSRLNCVFYCLSSSSALLALEALTFLG
                QLRGISPVPRTDIYITSSCLECVLETSVVPNQGETLNELLNHCHHLVERVPPEPI
                KGLFESELQNLGLKVHIATDFIEQSVGKHEAVLQESLAYLKAHTIFNNTPKQVLELSN
                LLYWNSGQNPSDSGVKCSELSKIWSRENELQKYRPKLNNGEPPGHFFDLHSPQGTEL
                LFCGGIPSSTHDTITALKQDCSNTFMKQTRLTGVAKRQNELFMRLSNILYGEVPTKP
                KQTESALKTCDQSDASKNQVLQEAELRKEAVLNKLSKEGFRKLQACLSTHEEMLNSQL
                SLKIWGSVVYKQSATLLNHFLFRQSWTQASLPPSVNGSPEQFENSKFIKSSLYVKSL
                SREYLSTLRLHFFALITGPLITQEGLFPSPPNVQLAHCLEAAHFMPHQKMLNEMIKP
                TMEPQDWICSNFNEFYTIHETDLNGVQYECWKYLRELVLSVALYNITWEKNLCIYRTD
                HSCPTACSSGIKEGLIVYTYESHAPLLILVYSNKKWIFDLYALLYAHMQLANNGAHR"

polyA_signal    15048..15053
                /evidence=not_experimental
                16751..19315
                /note="ORF08; similar to EBV BALF4, CMV UL55, HSV UL27"
                /codon_start=1
                /evidence=not_experimental
                /product="glycoprotein B"
                /protein_id="AAC58059.1"
                /db_xref="PID:g2337975"
                /db_xref="GI:2337975"
                /translation="MAHTGSTVCAFLIFAVLKNVFCQTPTSSSEVEDVIPEANTVSDN
                IIRQQRNNTAKGIHSDSAFPFRVCSASNIGDIFRFQTSHSCPNTKDEHNEGILLIF
                KENIVPVFRVYRKIVTTSIYNGIYADAVTNQHVFSKSVPIYETRRMDTIYQCYN
                SLDVTVGGNLLVYTDNDGSNMTVDLQPVDGLSNSVRRYHSQPEIHAEPGWLLGGYRRR
                TTVNCEVTETDARAVPPFRYFITNIGDTIEMSPFWSKAWNETEFSGGEPDRTLTVAKDY
                RVVDYKPRGTQPGLHTRIFVDKEEYLSWAQQFRNISYCRWAHWKSFDNAIKTEHGKS
                LHFVANDITASFYTPNTQTREVLGKHVCLNNTIESELKSRLAKVNDTHSPNGTAQYYL
                AEGTILLTQIQFAYDTLRAQINNVLEELSRAWCREQHRASLMWNELSKINPTSVMSSI
                YGRPVSAKRIGDVISVSHCVVVDQDSVSLHRSMRVPGRDKTHECYSRPPVTFKFINDS
                HLYKGQLGVVNEILLTTAVEICHENTEHYFQGGNNMYFYKNVRHVKTMPGVDVATLD
                TFMVLNLTLVENIDFQVIELYSREEKRMSTAFDIETMFREYNYYTQRVTGLRRDLTDL
                ATNRNQFVDAFGSLMDDLGVVGKTVLNAVSSVATLFSSIVSGIINFIKNPGGMLLFG
                LIAAVVTTVILLNRKAKRFAQNPVQMIYPDIKTITSQREELQVDPISKHELDRIMLAM
                HDYHASKQPESKQDEEQGSTTSGPADWLNKAKNVLRRAGYKPLKRTDSFESTGVP"

polyA_signal    19314..19319
                /evidence=not_experimental
                19428..22508
```

```
alignment_scores:
                    Quality:  1342.00    Length:    590
                      Ratio:     3.092     Gaps:       5
Percent Similarity:  73.559    Percent Identity:  45.932

alignment_block:
US-09-041-236-2 x AF005370

Align seg 1/1    to: AF005370    from: 1    to: 130608
```

```
   2  LeuLeuLeuTrpAlaAlaAlaAlaSerAlaAlaGlnGlyHisLeuArgSe  18
      |||||||:::::::||||||| ::::::::||:::||||
3660  TTACTGATGATTTATCAGCCATCACAGCTGCTAAATCTGGTTTATAGA  3709

  18  rGlyProArgIlePheAlaValTrpLysGlyHisValGlyValGlnAspArgV  35
      |||||             ||||||  |||||
3710  TAAGCCAAGGCTGAATTGTTAACCTAACTGATGGGTTGGACAGCACCGGT  3759

  35  alAspPheGlyGlnThrGluProHisPheHisValLeuProGly  51
      :: ::||||||  :::||||||||||| ::
3760  TT...TTTGGACCCCAGGAACCACCACTGTCTTTTCACAGCCTCAAC  3806

  52  SerSerValTrpValGlyGlyArgGlyLysValTyrLeuPheAspPh  68
      ||::::::||::||| ::||||||||
3807  TCTTCAGACGTATGTGGGAGGTAATAATACCATCTATTTGTTTGATTT  3856

  68  eProGluGlyLysAsnAlaSerValArgThrValAsnIleGlySerThrL  85
      :::::::|||||||::  ::||||| |||||
3857  TGCTCAGAGTCCAACGCCATCCACAGCTTTGATAAACATAACTAGCACAC  3906

  85  ysGlySerCysLeuAspLysArgAspCysGluAsnTyrIleThrLeuLeu  101
      ::||| ||||||||||| ||||||||
3907  ATAATACCCACCGGTTATCTAGTACCTGCGAAAACTTTATAACTCTGCTT  3956

 102  GluArgArgSerGluGlyLeuLeuAlaCysGlyThrAsnAlaArgHisPr  118
      ::::::::::||||||||||||||||||:::::::::
3957  CATAACCAGACAGATGGGCTGCTAGCTTGTGTACTAACTCACAGAAACC  4006

 118  oSerCysTrpAsnLeuValAsnGlyThrValValProLeuGlyGluMetA  135
      ||||||||   ||||||||| |||||||
4007  CAGCCGCTGGCTGGATAAAACACCTAACAACTCAAATTTTGGGGCCAAAC  4056

 135  rgGlyTyrAlaProPheSerProAspGluAsnSerLeuValLeuPheGlu  151
      ||||||   ||||||||||||||||||||
4057  TAGGCTTAGCCCCCTTCTCACCATCTGCCAATCTGGTCGTGTTTGAC  4106

 152  GlyAspGlyValTyrSerThrIleArgLysGlnGluTyrArgGlyLysIl  168
      ::|||||||||| ::  |||||||||||||||
4107  CAGAATGACACCTATTCCACCATTAACCCTCTACAAGAGCCTCAGTGGTC  4156

 168  eProArgPheArgIleArgGlyGluSerGluTyrThrSerAspI  185
      ::||||||||| |||::  ||||||||||||||||||
4157  TCACAAGTTTAGGAGGATCCGCTGGCCAAGTAGAACTATACACGAGTGACA  4206

 185  hrValMetGlnAsnProGlnPheLeuLysAlaThrIleValHisGlnAsp  201
      ||| :::||  ::::|||  ::||:::||||||
4207  CCGCCATGCAACCAGTTGCTCCAGGCAACAGCTGTGTCATAAAAAT  4256

 202  GlnAlaTyrAspLysIleTyrPheArgGlyGluAspAsnProAs  218
      :::|||||||||||||||||||||::::||  :::||
4257  GAATCTTATGATGATAAAATCTACTTTTCTTTCAAGAAAACAGCCACAG  4306

 218  pLysAsnProGluAlaProLeuAsnValSerArgValAlaGlnLeuCysA  235
      :  ::||||  ::||   ||||||::::||
4307  TGACTTCAAAACAGTTTCCACATACTGTACCTAGAGTGGGGCAGGTGTGCT  4356

 235  rgGlyAspGlyGlyGluSerSerValSerLysTrpTrpAspAsnThr  251
      ::||||||||||||||||||||  :::||
4357  CTAGTGATCAAGGTGGGAGAGCTCCCTGTCTGTCTCACAAGTGGACCACC  4406

 252  PheLeuLysAlaMetLeuValCysSerAspAlaAlaThrAsnLysAsnPh  268
      :::||||||||||||||||||||  :::|||
4407  TTTTTAAAAGCCAGACTGGTGTGTATACTGATGATACTGGAAGAATCTA  4456

 268  eAsnArgLeuGlnHisAspValPheLeuGluProAspSerProSerGlyInTrpA  285
      :||:::||||:::||  |||||| |||||||::
4457  CAATGAGCTACAAGATATTTTCATCTGGCAAGCCCCAGAGAACAGCTGGG  4506

 285  rgAspThrArgValTyrGlyValPheSerAsnProTrpAsnTyrSerAla  301
      ::::||||||::||||:::  ||||||||||||
4507  AAGAGACTCTCATCTATGGACTTTTTCGAGCCCCGGGACTTTTCTGCG  4556

 302  ValCysValTyrSerLeuGlyLeuGlyAspIleAspIlePheLysValPheArgSerSe  318
```

```
4557  GTCTGTGTGTTTACTGTGTAAAGGACATTGACCATGTGTTAAGACATCCAA  4606

 318  rLeuLysGlyTyrHisSerLeuProAsnProArgProGlyLysCysL  335
      ||||||:::|| |||::|||||||||||||
4607  GTTAAAAATATCATCATAAACTCCCACACCTAGACCAGGGCAATGCA  4656

 335  euProAspGlnGlnProIleProThrGlnThrPheGlnValAlaAlaAspArg  351
      :: ::|||||:::|| ||||||||||||||||||||
4657  TGAAGAACCATCAGCATGTCCCACAGAAACCTTCAGGTTGCTGACAGA  4706

 352  HisProGluValAlaGlnArgValGluProMetGlyProLeuLysThrPr  368
      :::||| ||||||||||||||||||||||::
4707  TATCCAAGAGTTGCAGATCCTGTATATCAGAGAACAATGCCATGTTCC  4756

 368  oLeuPheHisSerLysTyrHisTyrGlnLysValAlaAlaValHisArgMetG  385
      |::::|||||||||||||:::||||||||||:
4757  AATAATCAGTCAAATATATCATCACCACCAAACTACTTGTTTATAGGGTAG  4806

 385  lnAlaSerHisGlyGluThrPhe......HisValLeuTyrLeuThrThr  399
      ::||| |||  :::|||||||||||
4807  AG......TATGGAGGTGTTTTTGGCAACTATTTTTTTACCTCACTACC  4850

 400  AspArgGlyThrArgIleHisLysValValGluProGlyGlnGluHisSe  416
      ||||:: ||||||||||||||||||||||
4851  ATCAAAGGGACTATTCATATATATGTGGAGGTGAAGAATTCCAACTCTAC  4900

 416  rPheAlaPheAsnIleMetGluIleGlnProPheArgArgAlaAlaAlaI  433
      :::||||||:::||||||||||||
4901  AACAGCTCTCAACATTTTAGAGATAAATCCCTTTCAGAAGCCAGCCCCCA  4950

 433  leGlnThrMetSerLeuAspAlaGluArgArgLysLeuTyrValSerSer  449
      |||:::|| ||||   ::||||||||||||::||
4951  TACAGAATATTCTTTTAGATAAATACAAATCAAAGCTTTATGTAAATTCA  5000

 450  GlnTrpGluValSerGlnValProLeuAspLeuCysGluValTyrGlyGl  466
      ||||::||||||||| |||||||||||
5001  GAGTGGGAGGTGAGTGAGTGGCATTAGACCCATGTGTCAGTGTATGGGAA  5050

 466  yGlyCysHisGlyCysLeuMetSerGluArgArgAspProTyrCysTrpAspProG  483
      ::||| :::|||||||||||||||||
5051  TGATTGTTTCAGCTGTTTATGTCAAGGAACCATGTTCAGTGTCATGGTATA  5100

 483  lnGlyArgCysIleSerIleTyrArgSerGluArgSerValLeuGlnSer  499
      ::::|||   ::::||||||::||||
5101  ACAACACCTGT.......TCCTTTAAACAGAGAGTATCTGTTGAAACC  5141

 500  IleAsnProAlaGluPro......HisLysGluCysProAsnProLysPr  514
      :::||||| ::::||
5142  GGTGGGTCCAGCTAACCGCACCCTTCAGAAATGTGTGGTGACCACTATGC  5191

 514  oAspLysAlaProLeuGlnLysValSerLeuAlaProAsnSerArgArgTyrT  531
      :::|||||||||||||||||::
5192  TCCAACTGTGGTTAAGCATCAAGTTTCTATACCTCTATTATCTAATTCTT  5241

 531  yrLeuSerCysProMetGluSerArgHisAlaThrTyrSerTrpArgHis  547
      :: ::||||||||||:::||||
5242  ATTTGTCCTGCCCAGCAGTCTCAAACCACGCTGACTACTTTTGGACTAAA  5291

 548  LysGluAsnValGluGlnSerCysProValProGlyHisGlnSerProAsnCy  564
      :::|||||||||||||||||::
5292  GGTGGTTCACAGAAAAAGATGCCCATGTCAACAAAAATGACTG  5341

 564  sIleLeuPheIleGluLeuLeuThrAlaGlnThrGlyHisTyrPheC  581
      |||:::  ||||||||||:::||||
5342  CATCTTGCTTATAGCTAGCACAGCACGACGCCACTACAGGAACCCACGTGT  5391

 581  ysGluAlaGlnGlyGlySer  587
      ::||| :::||||||||
5392  GCAACATGAAAGAAGATTCG  5411
```

seq_name: gb_vi:AHU18243

```
seq_documentation_block:
LOCUS        AHU18243      11874 bp     DNA                VRL      29-SEP-1995
DEFINITION   Alcelaphine herpesvirus 1 putative semaphorin homolog (AHV-sema)
             and putative membrane antigen genes, complete cds, and major
             ssDNA-binding protein gene, partial cds.
ACCESSION    U18243
NID          g1000716
VERSION      U18243.1  GI:1000716
KEYWORDS     .
SOURCE       Alcelaphine herpesvirus 1.
ORGANISM     Alcelaphine herpesvirus 1
             Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
             Gammaherpesvirinae.
REFERENCE    1  (bases 1 to 11874)
AUTHORS      Ensser,A. and Fleckenstein,B.
TITLE        Alcelaphine herpesvirus type 1 has a semaphorin-like gene
JOURNAL      J. Gen. Virol. 76 (Pt 4), 1063-1067 (1995)
MEDLINE      97201573
REFERENCE    2  (bases 1 to 11874)
AUTHORS      Ensser,A.
TITLE        Direct Submission
JOURNAL      Submitted (07-DEC-1994) Armin Ensser, Institut fuer Virologie,
             Friedrich-Alexander-Universitaet Erlangen-Nuernberg, Schlossgarten
             4, Nuremberg, Bavaria 91054, Germany
FEATURES             Location/Qualifiers
     source          1..11874
                     /organism="Alcelaphine herpesvirus 1"
                     /strain="C500"
                     /specific_host="Connochaetes taurinus taurinus"
                     /db_xref="taxon:35252"
                     /lab_host="Bos taurus"
                     /clone="clones C238, C249, C500BS"
     repeat_region   <1..(330,346)
                     /note="gamma-2 herpesviral GC-rich terminal tandem
                     repeats"
                     /function="putative packaging signals"
                     /rpt_type=tandem
                     3573..5534
     CDS
                     /standard_name="AHV-sema"
                     /note="putative semaphorin homolog; similar to Genbank
                     Accession Number L26081"
                     /codon_start=1
                     /protein_id="AAC54475.1"
                     /db_xref="PID:g1000717"
                     /db_xref="GI:1000717"
                     /translation="MAYLNATVSKPVISLLSLSKKVLKFEHCGGGGQCLGLITEFVIH
                     PAAMGTLCVSIRLLMLSAITAAKSRFIDKPRLIVNLTDGFGQHRFFGPEPHTVLFH
                     SLNSSDVYVGGNNTIYLFDFAHSSNASTALINITSTHNTHRLSSTCENFITLLHNQTD
                     GLLACTNSQKPSCWLINNLTTQFLGPKLGLGLAPFSPSSGNLVLFDQNDYISTINLYKS
                     LSGSHKFRRIAGQVELYTSDTAMHRPQFVQATAVHKNESYDDKIYFFFQENSHSDFKQ
                     FPHTVPRVGQVCSSDQGGESSLSVYKWTTFLKARLACVDYDTGRIYNELQDIFTWQAP
                     ENSWEETLIYGLFLSPWNFSAVCCVFTVKDIDHVFKTSKLKNYHHKLPTPRPGQCMKNH
                     QHVPTETFQVADRYPEVADRQVSNNAMFPIIQSKYIYTKLLVYRVEYGGVFWATIFY
                     LTTIKGTIHIYVRYEDSNSTTALNLLEINPFQKPAPIQNILLDNTNLKLYVNSEWEVS
                     EVPLDLCSVYGNDCFSCFMSRDPLCTWYNNTCSFQRVSVETGGPANRTLSEMCGDHY
                     APTVVKHQVSIPLLSNSYLSCPAVSNHADYFWTKDGFTEKRCHVKTHKNDCILLIANS
                     TTATNGTHVCNMKEDSVTVKLLEVNVTLM"
     CDS             6414..10523
                     /note="putative membrane antigen; similar to p140 (PIR
                     Accession Number S29606), and p160 (Genbank Accession
                     Number 226584)"
                     /codon_start=1
                     /protein_id="AAC54476.1"
                     /db_xref="PID:g1000718"
                     /db_xref="GI:1000718"
                     /translation="MITRELQLLAVTAEPSALETAAIAALRSIPGLQNVIIQTQDAYL
                     VTFFTSPRPLREHQxKIETLLFIKAALRTFDEQQYLPIPLSALSSSFFTFVYGPDINRL
                     PTSQSNELTALQTRDARERKVFNVERIQHCRLLFFSGPGSEQLNHTHYALLREILCGD
                     LTLYQPFRNELFNSSLDFYATIYPPRFLRDKVRHYVHGDINLNITGAASNYNWASPHT
                     AVTDSEIFRIPYTHVVFRIDRWEIRVRVPDVGFHGSRGVRFRQHGGAAYMCSTLNSTT
                     GATGCLQKGMTSFSISQPTLGQLGLITEPMCGWGQQELSATAIOMLTKYANTVEAFAK
                     TLNYAGVPLVQGFVTLSPPILREKRAMAFTGSLSITGLPPYILKPPTAETLDAAMERNR
                     QLLLVEVGYPDYGKGKLHHPVNIMNSESGRIAHILHQALRQLMIITPFGKVIHICCDW
```

```
QYNEKVTMFKIAKACGDLGLSMSARSLPAHTLRLLKKWFRNLNYNHRIIKNSWLKVD
SAAALVVIADNEYQDVSVKKMDIALSGWGCPFHLGNLTPNSNTTIVISDKNQYGEIVDI
QYKMHMPKQPSEGTEDTPLAPDMSNIQLFKNLDVTEDLLQVLRHPTVGCKAHIVHHV
DRCGNGHIAQQPGVGPFDIPLCDFSVTVHNLVDGDIREGMESVPRVWAADWRVARRLI
ETQYSTPGLDITDATLANNLGTYHIPSEHQVTVESKKYGNCIGIGEKTTFTQRDPLL
GTILAIVESCTNCILGPVENYEEFLIGLSISVPEGIHYRHEVNSIMAMAKDFCSSMNF
GFQVNSAENGNCLLRSVVATANAPCVVPGPSLKPYFKKPGSAILRVNLHTEHFLSGGI
CCMASGIGASETFTPTPSQLRNLLQFMLIVKAENLALSGHDVSDGGLICAVCEMFAG
GLSARLIIHDEDEPVFPLFSETPGFVLEVNAIDVAAIIARANLYVECIQIGEVVES
DIFTVFHQNTQLLSVPVSRLKHNWTLFSKSVDLLYVKEDQVLPEETSYGNYEVHLTVD
PYSIISQSTTRPNVLVHLLPGCGYPDALLAALTNSGFSPDTVVYPGCKYRHNREDAA
PGSPTADDFIAGIVLYGSSNIDSDVGDSTIRQWLNVNRQVINDVRRNLKAKGSFTLAI
GQLACRILFATKAIGFDAGSQQTPFLLPNASRRYESRWLNFKIPEDTKAVAFRDLRGC
VLPCWVQGTHLGFSHNNITFFGDLETRQQVAATFNGPLVQSGPAREYPLNPTEAEHPY
AGLCSEDGRHLALLFDPCLAFNTWQWHQNQTGPGQGELPVSPWKLMFYRLYNWSKFHQ
HYRSLLRTNLRHTFNFETQQLDFPHADRRGAVPHDNPQYIPMDAMDPSQ"
   11341..>11874
   /codon_start=1
   /product="major ssDNA-binding protein"
   /protein_id="AAC54477.1"
   /db_xref="PID:g1000719"
   /db_xref="GI:1000719"
   /translation="MALKHQLNHAVEDNLGSKAPIGPCGFIYIYPETHFNFKEISLLG
DKYAEAGAFSLPLLHGVTVEEAFVPNVKAVYKKIDMTTVSVKLSTFYNRAIIFHNVEK
FESIFSGPGLGSLCKEACDLFGYVPFTPLGEGSTDVSDICPPVWQEKDAIMAVVITEG
FKERLHLGKLIYLKSQMH"
```

```
CDS

BASE COUNT     3437 a    2699 c    2362 g    3376 t
ORIGIN

alignment_scores:
                        Quality:  1342.00          Length:       590
                        Ratio:       3.092           Gaps:          5
Percent Similarity:  73.559      Percent Identity:  45.932

alignment_block:
US-09-041-236-2 x AHU18243

Align seg 1/1  to: AHU18243    from: 1  to: 11874

   2 LeuLeuLeuTrpAlaAlaAlaAlaSerAlaGlnGlyHisLeuArgSe 18
      |||||::|||||||||||::|||     :::|||||::|||||||| ::
3741 TTACTGATGATTTATCAGCCATCACAGCTGCTAAATCTCGGTTTATAGA 3790

  18 rGlyProArgIlePheAlaValTrpLysGlyHisValGlyGlnAspArgV 35
      ||||||||::|||||||||||||::||||||||     :::||| |||
3791 TAAGCCAAGGCTGATTGTTAACCTAACTGATGGGTTTGGACAGCACCGGT 3840

  35 aLAspPheGlyGlnThrGluProHisThrValLeuPheHisGluProGly 51
      |||:|     |||:::|||||::||||||     ||||||||||:::||
3841 TT...TTTGGACCCCAGGAACCACACTGTGCTTTTCACAGCCTCAAC 3887

  52 SerSerSerValTrpLysValGlyAspGlyLysValTyrLeuPheAspPh 68
      ||||||:::|||::|||     :::|||:::||||||||||||:::||||
3888 TCTTCAGACGTATATGTGGGAGGTAATAATACCATCTATTTGTTTGATTT 3937

  68 eProGluGlyLysAsnAlaSerValArgThrValAsnIleGlySerThrL 85
      :::|||||||||||||||     :::|||||||||||||     ||||||
3938 TGCTCACAGCTCCAACGCCATCACAGCTTTGATAAACATAACTAGCACAC 3987

  85 ysGlySerCysLeuAspLysArgAspCysGluAsnTyrIleThrLeuLeu 101
      ||||||     :::|||:::||||||||||||     :::|||||||||
3988 ATAATACCCACCGGTTATCTAGTACCTGCGAAAACTTTATAACTCTGCTT 4037

 102 GluArgArgSerGlyGlnValLeuAlaCysGlyThrAsnAlaArgHisPr 118
      :::|||     |||     :::|||||||||     :::|||||||||||
4038 CATAACCAGACAGCAATGGGCTGCTAGCTTGTGGTACTAACTCACAGAAACC 4087

 118 oSerCysTrpAsnLeuValAsnGlyThrValProLeuGlyLeuGlyLeuMeta 135
      |||||||||     ||| |||||     :::|||     :::|||||||
4088 CAGCTGCTGGCTGATAAACAACCTAACAACTCAATTTTGGGGCAAAAC 4137

 135 rgGlyTyrAlaProPheSerProAspGluAsnSerLeuValLeuPheGlu 151
```

```
4138    |||   ||||||||||||||||:::::|||||||||
        TAGGCTTAGCCCCCTTCTCACCATCATCTGGCAATCTGGTGCTGTTGAC  4187
152  GlyAspGluValTyrSerThrIleArgLysGlnGluTyrAsnGlyLysIl  168
     ::|||||||||||||||:::::  ::    |||    |||
     GAATCTTATGATGATAAAATCTACTTTTCTTTCAAGAAAACAGCCACAG  4188
```

```
5032    TACAGAATATTCTTTAGATAATACAAATCTAAAGCTTTATGTAAATTCA  5081

450  GlnTrpGluValSerGlnValProLeuAspLeuCysGluValTyrGlyGl  466
     :::||||||||||||||::||  ||||||:|||||||::|||||:|||
5082 GAGTGGGAGGTGAGTGAGGTGCCATTAGACCTGTCAGTGTATGGGAA  5131
```

seq_name: gb_pat:AR040598

seq_documentation_block:
```
LOCUS       AR040598                 2601 bp    DNA              PAT      29-SEP-1999
DEFINITION  Sequence 53' from patent US 5807826.
ACCESSION   AR040598
NID         g5959961
VERSION     AR040598.1    GI:5959961
KEYWORDS
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2601)
  AUTHORS   Goodman,C.S., Kolodkin,A.L., Matthes,D., Bentley,D.R. and
            O'Connor,T.
  TITLE     Semaphorin gene family
  JOURNAL   Patent: US 5807826-A 53 15-SEP-1998;
            US-09-041-236-A 53 15-SEP-1998;
  FEATURES  Location/Qualifiers
     source          1..2601
                     /organism="unknown"
BASE COUNT      809 a    533 c    593 g    666 t
ORIGIN
```

alignment_scores:
```
        Quality:   649.50              Length:       687
        Ratio:       1.620             Gaps:          27
        Percent Similarity: 58.370     Percent Identity: 28.093
```

alignment_block:
```
US-09-041-236-2 x AR040598

Align seg 1/1  to: AR040598    from: 1   to: 2601

     1  LeuLeuLeuLeuTrpAlaAlaAlaAlaSerAlaAlaGlnGlyHisLeuAr  17
```

```
907 CCAAATGGCCATTGACACTCATTTTGATGAACTGCAGGATGTATTCCTAAT 956
      :::::::||| :::::::||||||||::::::::||:::||||||
 34 ATTGTCTGTCTTTTCTGGGGAGTATTACTTACAGCAAGAGCAAACTATCA 83

277 u.......ProAspProSerGlyGlnTrpArgAspThrArgValTyrGlyv 292
 17 gSerGly..........ProArgIlePheAlaValThrLysGlyHisv 30

957 GAACTTTAAAGATCCT.......AAAATCCAGTGTATATGGAG 994
      :||||::            :::::||||::::||||||
 84 GAATGGGAAGAACAATGTGCCAAGGCTGAAATTATCCTACAAAGAAATGT 133

292 alPhe........SerAsnProTrpAsnTyrSerAlaValCysValTyr 305
 30 aIGlyGlnThrArgGlu..........PheGlyGlnThrGlu 41

995 TGTTTACGACTTCCAGTAACATTTCAAGGGATCAGCCGTGTGTATGTAT 1044
      ||||||:        :::::::::::||||||||::::::::
134 TGGAATCCAACAATGTGATCACTTTCAATGGCTGGCCAACAGCTCCAGT 183

306 SerLeuGlyAspIleAspLysValPhe.............. 314
 42 ProHisThrValLeuPheHisGluProGlySerSerValTrpValGl 58

1045 AGCATGAGTGATGTGAGAAGGGTGTTCCTTGGCCCATATGCCCACAGGA 1094
       :::||:|||:::||||| |||:::::::||::||||
184 TATCATACCTTCCTTTTGGATGAGGAA...CGGAGTAGGCTGTATGTTGG 230

315 ....ArgThrSerSerLeuLysGlyTyrHisSerSerLeuProAsnProA 330
 58 yGlyArgGlyLysValTyrLeuPheAspPheProGluGlyLysAsnAlas 75

1095 TGGACCCAACTATCAATGGGTGCCTTATCAAGGAGAGAGTCCCTATCCAC 1144
       ::::::|||||||:::::::||:::::||||||||::::::::
231 AGCAAAGGATCACATATTTCATTCGACCTGGTTAATATCAAGGATTTTC 280

330 rgProGlyLysCys................LeuProAspGlnGln 339
 75 erValArgThrValAsnIleGlySerThrLys.............. 85

1145 GGCCAGGAACTTGTCCCAGCAAAACATTGGTGGTTTTGACTCTACAAAG 1194
       ||||||:| ::: :| :::::||
281 AAAAGATTGTGTGGCCAGTATCTTACACCAGAGAGGATGAATGCAAGTGG 330

340 ProIleProThrGluThrPheGlnValAlaAlaAspArgHisProGluValAl 356
 86 ...GlySerCysLeuAspLysArgAspCysGlyGluAsnTyrIleThrLeuLe 101

1195 GACCTTCCTGATGTGTTATAACCTTGCAAGAGTCATCCAGCCATGTA 1244
       ::::||| |||  :::::||  ::::||||||||||
331 GCTGGAAAAGACATCCTGAAA...GAATGTGCTAATTTCATCCAAGGTACT 377

356 aGlnArgValGluProMetGly...............ProLeuLysThrProL 369
101 uGluArgArgSerGlu..GlyLeuLeuAlaCysGlyThrAsnAlaArgH 117

1245 CAATCCAGTGTTTCCTATGAACAATCGCCCAATAGTGATCAAAACGGATG 1294
       |||::||| :::::|:::||||||:::::||||| |
378 TAAGGCATATAATCAGACTCACTTGTACGCCTGTGGAACGGGGGCTTTTC 427

369 euPheHisSerLysTyrHisTyrGlnLysValAlaValHisValAlaValHisArgMetGln 385
117 isProSerCysTrpAsnLeuValAsnGlyThrValValProLeuGlyGlu 133

1295 TA.......AATTATCAATTTACACAAATTGTCGTAGACCGAGTGGAT 1335
       ||       ::::||||||   :::::|||
428 ATCCAATTTGC...............ACCTACATTGAAATTGGACAT 459

386, AlaSerHisGlyGluThrPheHisValLeuTyrLeuThrHisAspArgG1 402
134 Met........................ 134

1336 GCAGAAGATGGACAG...TATGATGTTATGTTTATCGGAACAGACGTGTGG 1382
       ||||::|| ::::||   :::::||
460 CATCCTGAGGACAACAATATTTTTAAGCTGGAGAACTCACATTTGAAAACGG 509

402 yThrIleHisLysValValGlu...ProGlyGluGlnGluHisSerPheA 418
135 ArgGlyTyrAlaProPheSerProAspGlnSerLeuValLeuPheG 151

1383 GACCGTTCTTAAAGTAGTTTCAATTCCTAAGGGAGACTTGGTATGATTTAG 1432
       ::::||| |||  ||||:::::|||
510 CCGTGGGAAGAGTCCAATGACCCTAAGCTGCTGACAGCAGCATCCCCTTTAA 559

418 laPheAsnIleMet...GluIleGlnProPheArgArgAlaAlaAlaAlaIle 433
151 luGlyAspTyrSerThrIleArgLysLysGlnGluTyrAsnGlyLys 167

1433 AAGAGGTTCTGCTGGAAGAAAATGACAGTTTTCGGGAACCGACTGCTATT 1482
       ||:::||||  ::::|||||||:::||||
560 TAGATGGAGAATTATACTCTGGAACTGCAGCTGATTTTATGGGGGCGAGAC 609

434 GlnThrMetSerLeuAspAlaGluArgArgLysLeuTyrValSerSerSerGl 450
168 IleProArgPheArgIleArgGlyIleArgGlyGluSerGluLeuTyrThr...... 182

1483 TCAGCAATGGAGCTTTCCACTAAGCAGCAACAACTATATATTGGTTCAAC 1532
       :::|| :::||||  :::::|||||
610 TTTGCTATCTTCCGAACTCTGGGCACCACCAATCAGGACAGCAGCA 659

450 nTrpGluValSerGlnValProLeuAspLeuCysSerGluValTyrGlyGlyG 467
183 ....SerAspThrValMetGlnAsnProGlnPheIleLysAlaThrIlev 198

1533 GGCTGGGGTGTTGCCTCCAGCCTCCCTTTACACCGGTGATATTTACGGGAAAG 1582
       ::::|||| |||:::::||||||
660 GCATGAATTCCAGGTGGCTCAATGATCCAAAGTTCATTAGTAGTGCCCACCTCA 709

467 lyCysHisGlyGlyLeuMetSerArgAspProTyrCysSerGlyTrpAspGln 483
198 alAlaGlnAsnPro...AspAspThrAspPheLysAspAlaTyrPhe 213

1583 CGTGTGCTGAGTTGGCCTCGCTATTTCCCCGAGACCCTTACTGTGCTGGGATGGT 1632
       ::::|||| :::::|||||||
710 TCTCAGAGAGTGACAATCCTGAAGATGACAAAGTTACTTTTTCCGT 759

484 GlyLysGluArgSerValLeuGl......SerSerGluValLeuGl 498
214 GluAspAsnProGluGlnPheAsnProLeuGlyGlyGluSerLeuSerVal 230

1633 TCTGCAGTTCTCGCTATTTCGCAAAGAGACGCCACAAGACGACA 1682
       :::::||||  :::::|||||
760 GAAAATCAATAGATGGACAATAGGAGAACACTCGGAAAAGCTACTCACGCTGAAT 809

498 nSerIleAsnProAlaGluProHisLysGluCysProAsnProLysProA 515
230 lAlaGlnLeuCysArgGlyAspAsnProGlnGlnGlyLeuSerSerLeuSerVal 247

1683 AGTATAAGAAATGGAGACCCACCTGCACTGTTCAGACTTACACCATG 1732
       :::::|||| :::||| :::
810 AGGTCAGAGATGCAAGAATGACTTTGGAGGGCCACAGAAGTCTG...GTGA 856

515 spLys................AlaProLeuGlnLysValSerLeu...AlaPro 526
247 erLysTrpAspThrPheLeuLysAlaMetLeuValCysSerAspAlaAla 263

1733 ATAATCACCATGGCCACCAGCCCTGAAGAGAGAATCATCTATGGTGTAGAG 1782
       :|||||  ::::||||||  :::::|||
857 ATAAATGGACAACATTCCTCAAAGCTCGTCGTTGCTCAGTGCCAGGT 906

527 AsnSerArgTyrTyrLeuSerCysProMetGluSerArgHisAlaThrTy 543
264 ThrAsn........LysAsnPheAsnArgLeuGlnLeuAspAspValPheLeuLe 277

1783 AATAGTAGCACATTTTGAATGCAGTCCGAGTCGGCAGAGAGCCGCTGGT 1832
       :::|||||||||||||||||
277 .......(continuation)
```

```
543 rSerTrpArgHisLysGluAsnValGluGlnSerCysGluProGlyHisG 560     231 AGCAAAGGATCACATATTTCATTCGACCTGTGTTAATATCAAGGATTTTC 280
    ::::::       ::::::::      |||    |||     :::          erValArgThrValAsnIleGlySerThrLys.................
1833 CTATTGGCAAATTCCAGAGGCGCAAATGAAGAGCGAAAAGAAGAGATCAGAG 1882 75 erValArgThrValAsnIleGlySerThrLys................. 85
                                                           281 AAAAGATTGTGTGGCCAGTATCTTACACCAGAGAGATGAATCCAAGTGG 330
560 lnSerProAsnCysIle.............LeuPheIleGluAsnLeu 571     86 ...GlySerCysLeuAspLysArgAspCysGluAsnTyrIleThrLeuLe 101
    ::: ||        |||             ||||||::::::::::::||        ::: :::||| ||||||| :::|| |||    ||  :::||
1883 TGGATGATCATATCATCAGGACAGATCAAGGCCTTCGCCTACGTAGTCTA 1932 331 GCTGGAAAAGACATCCTGAAA...GAATGTGCTAATTTCATCAAGGTACT 377

572 ThrAlaGlnGlnTyrGlyHisTyrPheCysGluAlaGlnGluGlySerTy 588    101 uGluArgArgSerGlu...GlyLeuLeuAlaCysGlyThrAsnAlaArgH 117
    :::::::||||:::   |||||||||:::::||||| |||||             |:::  :::::        |||  ||||||||:::::::||| |
1933 CAACAGAAGGATTCAGGCAATTACCTCTGCCATGCGGTGGAACATGGGTT 1982 378 TAAGGCATATAATCAGACTCACTTGTACGCCTGTGGAACGGGGGCTTTTC 427

588 rPheArgGluAlaGlnHisTrpGlnLeuLeuProGluAspGlyIleMetA 605    117 lsProSerCysTrpAsnLeuValAsnGlyThrValProLeuGlyGlu 133
    ::::::         |||              :::  |||  |||            ||||| ||:
1983 CATCCAAACTCTTCTTAAGGTAACCCTG........GAAGTCATTGACA 2023 428 ACCTACATTGAAAATGGACAT....................ACCCAATTGC 459

605 laGluHisLeu 608                                         134 Met................................................... 134
    ::||||||||
2024 CAGAGCATTTG 2034                                        460 CATCCTGAGGACAATATTTTTAAGCTGGAGAACTCACATTTTGAAAACGG 509

                                                           135 .ArgGlyTyrAlaProPheSerProAspGluAsnSerLeuValLeuPheG 151
seq_name: gb_pat:I47054                                          |||||    :::|||||:::||              ::  ||||:
                                                           510 CCGTGGGAAGTCCATATGACCCTAAGCTGCTGACAGCATCCCTTTTAA 559
seq_documentation_block:
LOCUS       I47054     2601 bp    DNA          PAT       151 luGlyAspGluValTyrSerThrIleArgLysGlnGluTyrAsnGlyLys 167
DEFINITION  Sequence 53 from patent US 5639856.     07-OCT-1997   :::     :::
ACCESSION   I47054                                         560 TAGATGGAGAATTATACTCTGAACTCAGCTGATTTTATGGGGCGAGAC 609
NID         g2471019
VERSION     I47054.1  GI:2471019                          168 IleProArgPheArgIleArgGlyGluSerGluLeuTyrThr......... 182
KEYWORDS    .                                                  ||||:::    ||||||       |||  ::
SOURCE      Unknown.                                       610 TTTGCTATCTTCCGAACTCTTGGCACCACCAATCAGGACAGAGCA 659
ORGANISM    Unknown.
            Unclassified.                                 183 ....SerAspThrValMetGlnAsnProGlnPheIleLysAlaThrIlev 198
REFERENCE   1  (bases 1 to 2601)                               ::::::    :::::::||      ::    ::  :::
AUTHORS     Goodman,C.S., Kolodkin,A.L., Matthes,D., Bentley,D.R. and   660 GCATGATTCCAGGTGCTCAATGATCCAAAGTTCATTAGTGCCACCTCA 709
            O'Connor,T.
TITLE       Semaphorin gene family                        198 alHisGln...AspGlnAlaTyrAspLysIleTyrTyrProPheArg 213
JOURNAL     Patent: US 5639856-A 53  17-JUN-1997;             :::||||      |||    :::  ::  ::::::::||||||:::
FEATURES    Location/Qualifiers                            710 TCTCAGAGAGTGACAATCCTGAGATGACAAAGTATACTTTTTCTTCCGT 759
source      1..2601
            /organism="unknown"                           214 GluAspAsnProLysAsnProGluAlaProLeuAsnValSerArgVa 230
BASE COUNT  809 a    533 c    593 g    666 t                  |||                 ::|||            :::::
ORIGIN                                                     760 GAAAATGCAATAGATGGAGAACACCTGGAAAAGCTACTCACGCTAGAAT 809

alignment_scores:                                         230 lAlaGlnLeuCysArgGlyAspGlyAspGlnGlyGlySerLeuSerValS 247
    Quality:     649.50           Length:     687              ::::||:||::::|||::: |||||::     |||  :::
    Ratio:       1.620            Gaps:        27          810 AGGTCAGATATGCAAGAATGACTTTGGAGGCCACAGAAGTCTG...GTGA 856
Percent Similarity:  58.370      Percent Identity:  28.093
                                                          247 erLysTrpAsnThrPheLeuLysAlaMetLeuValCysSerAspAlaAla 263
alignment_block:                                               ::||||||  :::::::||      :::|||  ||||||||
US-09-041-236-2 x I47054                                  857 ATAAATGGACACACATTCCTCAAAGCTCGTCTGATTTGCTCAGTGCCAGGT 906

Align seq 1/1  to: I47054    from: 1 to: 2601             264 ThrAsn........LysAsnPheAsnArgLeuGlnAsnValPheLeuLe 277
                                                              :::||||||            |||:::|| ||||||:::||||||
  1 LeuLeuLeuLeuTrpAlaAlaAlaAlaSerAlaGlnGlyHisLeuAr 17    907 CCAAATGGCATTGACACTCATTTTGATGAACTGCAGGATGTATTCCTAAT 956
    ::::: |||::||::    ::|||||||::  ::|||::::: ::
 34 ATTGTCTGCTCTTTCTGGGGAGTATTACTTACCAGCAAGAGCCAAACTATCA 83   277 u......ProAspProSerGlyGlnTrpArgAspThrArgValTyrGlyYv 292
                                                              ||      :::            ||||||||||
 17 gSerGly..........ProArgIlePheAlaValTrpLysGlyHisV 30   957 GAACTTTAAAGATCCT.........AAAAATCCAGTTGTATATGGAG 994
    ::|||||          |||||| ||| :::||::|||:::
 84 GAATGGGAAGAACAATGTGCCAAGGCTGAAATTATCCTACAACAAAAATGT 133  292 alPhe........SerAsnProTrpAsnTyrSerAlaValCysValTyr 305
                                                              ||||||        :::|||::||||::::::::::||
 30 alGlyGlnAspArgValAsp.............PheGlyGlnThrGlu 41   995 TGTTTACGACTTCCAGTAACATTTCAGGGATCAGCCGTGTGTATGTAT 1044
    ::  :::::::||||::
134 TGGAATCCAACAATGTGATCACTTTCAAATGGCTTGGCCAACAGCTCCAGT 183  306 SerLeuGlyAspIleAspLysValPhe......................... 314
                                                              ||:::  :::      :::
 42 ProHisThrValLeuPheHisGluProGlySerSerValTrpValGl 58   1045 AGCATGAGTGATGTGAGAAGGGGTTCCTTGGTCCATATGCCCACAGGGA 1094
    ||||||  |||||:::  ::: ||| :::  :::||||:::|||||
184 TATCATACCTTCCTTTGGATGAGGAA...CGGAGTAGGCTGTATGTTGG 230  315 ....ArgThrSerLeuLysGlyTyrHisSerSerLeuProAsnProA 330
                                                              ::      |||:::
 58 yGlyArgGlyLysValTyrLeuPheAspPheProGluGlyLysAsnAlaS 75  1095 TGGACCCAACTATCAATGGGTGCCTTATCAAGGAAGAGTCCCCTATCCAC 1144
    |:::::   :::::::|||:::    ||:::
```

605 laGluHisLeu 608
    ::||||||||||
2024 CAGAGCATTTG 2034

seq_name: gb_pr2:HUMHSEM

seq_documentation_block:
LOCUS   HUMHSEM   2530 bp   mRNA   PRI   08-MAY-1995
DEFINITION Homo sapiens semaphorin-III (Hsema-I) mRNA, complete cds.
ACCESSION   L26081
NID   g799328
VERSION   L26081.1 GI:799328
KEYWORDS   semaphorin.
SOURCE   Homo sapiens (human).
ORIGIN   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 2530)
AUTHORS   Kolodkin,A.L., Matthes,D.J. and Goodman,C.S.
TITLE   The semaphorin genes encode a family of transmembrane and secreted
            growth cone guidance molecules
JOURNAL   Cell 75 (7), 1389-1399 (1993)
MEDLINE   94094332
COMMENT   On May 8, 1995 this sequence version replaced gi:436559.
FEATURES   Location/Qualifiers
     source   1..2530
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /dev_stage="foetus"
                     /tissue_type="brain"
     gene   16..2331
                     /gene="Hsema-III"
     CDS   16..2331
                     /gene="Hsema-III"
                     /codon_start=1
                     /product="semaphorin-III"
                     /protein_id="AAA65938.1"
                     /db_xref="PID:g436560"
                     /db_xref="GI:436560"
                     /translation="MGWLTRIVCLFWGVLLTARANYQNGKNNVPRLKLSYKEMLESNN
                     VITFNGLANSSSYHTFLLDEERSRLYGAKDHIFSFDLVNIKDFQKIVWPVSYTRRDE
                     CKWAGKDILKECANFIKVLKAYNQTHLYACGTGAFHPICTYIEIGHPEDNIFKLENS
                     HFENGRGKSPYDPKLLTASLLIDGELYSGTRADFMGRDFAIFRTLGHHHPIRTEQHDS
                     RWLNDPKFISAHLISESDNPEDKVYFFRENAIDGEHSGKATHARIGQICKNDFGGH
                     RSLVNKWTTFLKARLICSVPGPNGIDTHFDELQDVFLMNFKDPKNPVVYGVFTTSSNI
                     FKGSAVCMYSMSDVRRVFLGPYAHRDGPNYQWVPYQGRVPYPRPGTCPSKTFGGFDST
                     KDLPDDVITFARSHPAMYNPVFPMNNRPIVIKTDVNYQFTQIVVDRVDAEDGQYDVMF
                     IGTDVGTVLKVVSIPKETWYDLEEVLLEEMTVFREPTAISAMELSTKQQQLYIGSTAG
                     VAQLPLHRCDIYGKACAECCLARDPYCAWDGSACSRYFPTAKRRTRQDIRNGDPLTH
                     CSDLHHDNHHGHSPEERIIYGVENSSTFLECSPKSQRALVYWQFQRRNEERKEEIRVD
                     DHIIRTDQGLLRSLQQKDSGNYLCHAVEHGFIQTLLKVTLEVIDTEHLEELLHKDDD
                     DGSKTKEMSNSMTPSQKVWYRDFMQLINHPNLNTMDEFCEQVWKRDKQRQRPGHT
                     PGNSNKWKHLQENKKGRNRRTHEFERAPRSV"

BASE COUNT   786 a   518 c   576 g   650 t
ORIGIN

alignment_scores:
                    Quality: 649.50   Length: 687
                    Ratio: 1.620   Gaps: 27
Percent Similarity: 58.370   Percent Identity: 28.093

alignment_block:
US-09-041-236-2 x HUMHSEM

Align seg 1/1 to: HUMHSEM from: 1 to: 2530

    1 LeuLeuLeuLeuTrpAlaAlaAlaAlaSerAlaGlnGlyHisLeuAr 17
      :::::: |||||||:|||||:::::::::::::: ||::||||||::
   34 ATTGTCGTCGTTTTCTGGGGAGTATTACTTACAGCCAAGAGCAAACTATCA 83

   17 gSerGly..........ProArgIlePheAlaValTrpLysGlyHisV 30
      :::|||| ||||||::

330 rgProGlyLysCys................LeuProAspGlnGln 339
    ||||||||| ||| :: :: :: |
1145 GGCCAGGAACTTGTCCCAGCAAAACATTTGGTGGTTTTGACTCTACAAAG 1194

340 ProIleProThrGluThrPheGlnValAlaAspArgHisProGluValAl 356
    :::||| ::::::::: ||| :::::: ||||| :::
1195 GACCTTCCTGATGTGTTATAACCTTTGCAAGAGTCATCCAGCCATGTA 1244

356 aGlnArgValGluProMetGly.........ProLeuLysThrProL 369
    :::||| | |||:||||| :: |:: ||||||| :
1245 CAATCCAGTGTTTCCTATGAACAATCGCCCAATAGTGATCAAAACGGATG 1294

369 euPheHisSerLysTyrHisTyrGlnLeuValAlaValHisArgMetGln 385
    ::   ::::|||||::::||| |||:::::::|
1295 TA.......AATTATCAATTACACAAATGTCGTAGACCCAGTGGAT 1335

386 AlaSerHisGlyGluThrPheHisValLeuTyrLeuThrThrAspArgGl 402
    ||::: ::::: |||::::::::::::: |||||||| ||
1336 GCAAGAAGATGGACAG...TATGATGTTATGTTTATCGGAACAGATGTTGG 1382

402 yThrIleHisLysValGlu...ProGlyGlnGluHisSerPheA 418
    ||||| :|||:::||||::: ||:||||::::::::||||||
1383 GACCGTTCTAAAAGTAGTTCAATTCCTAAGGAGACTTGGTATGATTTAG 1432

418 laPheAsnLeuMet...GluIleGlnProPheArgArgAlaAlaAlaIle 433
    ::|||:::||| |||:::|||:::: ::::::||||||||
1433 AAGAGGTTCTGCTGGAAGAAATGACAGTTTTCCGGGAACCGACTGCTATT 1482

434 GlnThrMetSerLeuAspAlaGluArgLysLeuTyrValSerSerGl 450
    ::::: ||||:::||||||||||||:::|||:||||:||||
1483 TCAGACAATGGAGCTTTCCACTAGCAGCAACAACTATATTGGTTCAAC 1532

450 nTrpGluValSerGlnValProLeuAspAspTyrCysGlyTyrGlyGlyG 467
    :|||:::|||::::||||:::::|||||||::|||||| ||:||
1533 GGCTGGGGTTGCCCAGCTCCCTTTACACCGGTGTGATATTTACGGGAAAG 1582

467 lyCysHisGlyLysLeuMetSerArgAspProTyrCysGlyTrpAspGln 483
    ::||| ||| ::::::|||||:::::::||| :||||||||| |
1583 CGTGTGCTGAGTCTGCCTCGCCCGAGCACCTACTGTGCTCTGGGATGGT 1632

484 GlyArgCysIleSerIleTyr.....SerSerGluArgSerValLeuGl 498
    |||:::||||:::|||||| :::|||||:::|||:: ||||:||
1633 TCTCGCATGTTCTGCTATTTCCCACTGCAAAGACAGAGACCACAAGACA 1682

498 nSerIleAsnProAlaGluProHisLysGluCysProAsnProLysProA 515
    ::||||::::||||||:::||||:::|||||||:::::||:||||:||
1683 AGATATAGAAATGGAGACCCACTGACTCACTGTTCAGACTTACACCATG 1732

515 spLys.......... AlaProLeuGlnLysValSerLeu...AlaPro 526
    ||:::   ::: |||:||:::   :::::|| :::
1733 ATAATCACCATGGCCACAGCCCTGAAGAGATCATCTATGTGTAGAG 1782

527 AsnSerArgTyrTyrLeuSerCysProMetGluSerArgHisAlaThrTy 543
    :::|||::: ||||:::::|:|||:::||||||:::|||:|||||:|||
1783 AATAGTAGCACATTTTGGAATGCAGTCGGAGACGCGCTGGT 1832

543 rSerTrpArgHisLysGluAsnValGluGlnSerCysGluProGlyHisG 560
    ||||||||||||:::||||||:::|||:::||||||||||||||||:||
1833 CTATTGGCAATTCCAGAGGGCGAAATGAAGAGCGAAAAGAAGAGATCAGAG 1882

560 lnSerProAsnCysIle.............LeuPheIleGluAsnLeu 571
    ::|||||||||:::||| :::|||||||:::|||
1883 TGGATGATCATATCATCACAGACGATCAAGGCCTTCTGCTACGTAGTCTA 1932

572 ThrAlaGlnGlnTyrGlyHisTyrPheCysGluAlaGlnGlyGlySerTy 588
    |||:::::::::: |||:::::::|||||||:::|||:|||||||||:::
1933 CAACAGAAGGATTCAGGCAATTACCTCTGCCATGCGGTGGAACATGGGTT 1982

588 rPheArgGluAlaGlnIleTrpGlnLeuLeuProGluAspArgIleMetA 605
    :::||||:::::::::|:||||::::||||||:::|:::|||::||||
1983 CATACAAACTCTTCTTAAGGTAACCCTG.........GAAGTCATTGACA 2023

605 laGluHisLeu 608
    ::::||||
2024 CAGAGCATTTG 2034

```
 84  GAATGGGAAGAACAATGTGCCAAXGCTGAAATTATCCTACAAAGAAATGT  133

 30  aIGlyGlnAspArgValAsp.............PheGlyGlnThrGlu   41
     .:   :::::: |||||
134  TGGAATCCAACAATGTGATCACTTTCAATGGCTTGGCCAACAGCTCCAGT  183

 42  ProHisThrValLeuPheHisGluProGlySerSerValTrpValGl    58
     |||||| ||| |:|:||
184  TATCATACCTTCCTTTGATGAGGAA...CGGAGTAGGCTGATGTTGG      230

 58  yGlyArgGlyLysValTyrLeuPheAspPheProGluGlyLysAsnAlaS   75
     |:||:| ::::::: ||||| :  |||:::  :
231  AGCAAAGGATCACATATTTCAATCGACCTGGTTAATATCAAGGATTTTC   280

 75  erValArgThrValAsnIleGlySerThrLys................   85
     :: |:::::
281  AAAGATTGTGTGGCCAGTATCTTACACCAGAGAGATGAATGCAAGTGG   330

 86  ...GlySerCysLeuAspLysArgAspCysGluAsnTyrIleThrLeuLe  101
     |||:|||: ::  :||||| |||:::||| ::::|||
331  GCTGGAAAAGACATCCTGAAA...GAATGTGCTAATTTCATCAAGGTACT  377

101  uGluArgArgSerGlu...GlyLeuLeuAlaCysGlyThrAsnAlaArgH  117
     |:::  :::::: ||  |:||||||||||:::||| |
378  TAAGGCATATAATCAGACTCACTTGTACGCCCTGTGGAACGGGGCTTTTC  427

117  lsProSerCysTrpAsnLeuValAsnGlyThrValProLeuGlyGlu    133
     ||||| :: ::::  :||||| ::::::
428  ...ACCTACATTGAAATTGGACAT                           459

134  Met...............                                 134
     |||
460  CATCCTGAGGACACATATTTTAAGCTGGAGAACTCACATTTGAAAACGG  509

135  .ArgGlyTyrAlaProPheSerProAspGluAsnSerLeuValLeuPheG  151
     |||||| ::|||:::|||
510  CCGTGGAGAGAGTCCATATGACCCTAAGCTGCTGACAGCATCCCTTTTAA  559

151  luGlyAspValTyrSerThrIleArgLysGlnGluTyrAsnGlyLys    167
     ||||||:::|||:::|:: |||:||||||:||
560  TAGATGGAGAATTATACTCTGGAACTGCAGCTGATTTTATGGGGCGAGAC  609

168  IleProArgPheArgArgIleArgGlyGluSerGluLeuTyrThr....   182
     :::|||||||||| ::  ::
610  TTTGCTATCTTCCGAACTCTTGGGCACCACCAATCAGGACAGAGCA    659

183  ....SerAspThrValMetGlnAsnProGlnPheIleLysAlaThrIlev  198
     ::||:::|||::::||| ::::||| |||
660  GCATGAATTCCAGGTGGCTCAATGATCCAAAGTTCATTAGTGCCCACCTCA  709

198  alHisGln...AspGlnAlaTyrAspAspLysIleTyrTyrPhePheArg  213
     ::   ||| |||||||:::||| ||||||
710  TCTCAGAGAGTGACAATCCTGAAGATGACAAAGTATACTTTTTCTTCCGT  759

214  GluAspAsnProAspLysAsnProGluAlaProLeuAsnValSerArgVa  230
     |||||| ::  :: :||||:
760  GAAAATGCAATAGATGGAGAACACTCTGAAAAGCTACTCACCGTAGAAT  809

230  lAlaGlnLeuCysArgGlyArgGlnGlyGlyGluSerLeuSerValS    247
     :::|:||::: |||||: ||||| :||:
810  AGGTCAGATATGCAAGAATGACTTGGAGGGCACAGAAGTCTG...GTGA  856

247  erLysTrpAsnThrPheLeuLysAlaMetLeuValCysSerAspAlaAla  263
     ::|||||||||||::||||| |||| ||||:::
857  ATAAATGGACAACATTCCTCAAAGCTCCTGATTTGCTCAGTGCCAAGT   906

264  ThrAsn.......LysAsnPheAsnArgLeuGlnAspValPheLeuLe  277
     |||   ::: :|:::||||||||||:::|||
907  CCAAATGGCATTGACACTCATTTTGATGAACTGCAGGATGTAATTCCTAAT  956

277  u......ProAspProSerGlyGlnTrpArgAspThrArgValTyrGlyV  292
     :      |||||||
957  GAACTTTAAAGATCCT.........AAAAATCCAGTTGTATATGGAG  994
```

```
292  alPhe.................SerAsnProTrpAsnTyrSerAlaValCysValTyr  305
     |||||           :::: ||  |||||||||:::||
995  TGTTTACGACTTCCAGTAACATTTCAAGGATCAGCCGTGTGTATGTAT  1044

306  SerLeuGlyAspIleAspLysValGlyValPhe....................  314
     |||||||::|::||::||:||| |
1045 AGCATGAGTGATGTGAGAAGGGTGTTCCTTGGTCCATATGCCACAGGA  1094

315  ....ArgThrSerSerLeuLysGlyTyrHisSerGluProAsnProA     330
     :::   ::  |||||||:::||||| ::::||||
1095 TGGACCCAACTATCAATGGGTGCCTATCAAGGAGAGTCCCTATCCAC   1144

330  rgProGlyLysCys......................LeuProAspGlnGln  339
     |||||||| |||                        ::   ::
1145 GGCCAGGAACTTGTCCCAGCAAAACATTTGGTGGTTTGACTCTACAAAG  1194

340  ProIleProThrGluThrPheGlnValAlaAspArgHisProGluValAl  356
     ::|||                       ::::|||
1195 GACCTTCCTGATGATGTTATAACCTTGCAGAAGTCATCCAGCCATGTA  1244

356  aGlnArgValGluProMetGly.........ProLeuLysThrProL    369
     ::: |||||||           :: ||||||||||
1245 CAATCCAGTGTTTCCTATGACAATCGCCCATAGTGATCAAAACGGATG  1294

369  euPheHisSerLysTyrHisTyrGlnLysValAlaValHisArgMetGlu  385
     ::  |||||||||||:::||| |||||| ||:::|||
1295 TA............AATTATCAATTACCAAATTGCTGAGACCGAGTGGAT  1335

386  AlaSerHisGlyGluThrPheHisValLeuThrLeuThrThrAspArgGl  402
     |||||||||   |||                :::||||:::|||
1336 GCAGAGCATGGACAG...TATGATGTTATGTTTATCGGAACAGATGTTGG  1382

402  yThrIleHisLysValValGlu...ProGlyGluGlnHisSerPheA    418
     |||:::||||||          |||  ::|||
1383 GACCGTTCTTAAAGTAGTTTCAATTCCTAAGGAGACTTGGTATGATTTAG  1432

418  laPheAsnIleMet...GluIleGlnProPheArgArgAlaAlaAlaIle  433
     |:::|||        ::::|||::||||||| ::::||||
1433 AAGAGGTTCGTGCTGGAAGAAATGCAGTTTTCGGGAACCGACTGCTATT  1482

434  GlnThrMetSerLeuAspAlaGluArgArgLysLeuTyrValSerSerGl  450
     |||  ||||||||||||::::|||
1483 TCAGCAATGGAGCTTTCCACTGACAGCAACAACTATATATTGGTTCAAC  1532

450  nTrpGluValSerGlnValProLeuAspLeuCysGluValTyrGlyGlyG  467
     |||||| :::|||                  :::||||
/553 GGCTGGGGTTGCCCAGCTCCCTTTACACCCGTGTGATATTTACGGGAAAG  1582

467  lyCysHisGlyCysLeuMetSerArgAspProTyrCysGlyTrpAspGln  483
     ::||| |||                    ::::||||
1583 CGTGTGCTGAGTGTGCCTGCCCTGCCCTTACTGTGCTTGGGATGGT     1632

484  GlyArgCysIleSerIleTyr......SerSerGluArgSerValLeuGl  498
     |||                        ::   ::::|||
1633 TCTGCATGTTCTCGCTATTTCCCACTGCAAGAGCAAAGAGACAGACA     1682

498  nSerIleAsnProAlaGluProHisLysGluCysProAsnProLysProA  515
     ::::|||                      ::::|||
1683 AGAATAAAGAAATGGAGACCCACTGACTCACTGTTCAGACTTACACCATG  1732

515  spLys....................AlaProLeuGlnLysValSerLeu...AlaPro  526
     ||:                      :::
1733 ATAATCACCATGGCCACAGCCCTGAAGACGAAGAATCATCTATGGTGTAGAG  1782

527  AsnSerArgTyrTyrLeuSerCysProMetGluSerArgHisAlaThrTy  543
     ::::|||                      :::||||
1783 AATAGTAGCACATTTTGGATGACGTCCGAGAGTCCGAGATGCCTGGT  1832

543  rSerTrpArgHisLysGluAsnValGluGlnSerCysGluProGlyHisG  560
     |||                          ||| ||||||
1933 CTATTGGCAATTCCAGAGCGCGAAATGAAGAGCGAAAAGAAGAGATCAGAG  1882
```

```
 560 lnSerProAsnCysIle............LeuPheIleGluAsnLeu 571
     :::    :::   :::             |||:::::::::::|||
1883 TGGATGATCATATCATCAGGACAGATCAAGGCCTTCTGCTACGTAGTCTA 1932

 572 ThrAlaGlnInTyrGlyHisTyrPheCysGluAlaGlnGluGlySerTy 588
     :::::: |||:::||:::||:::||||::||| ||| :::::
1933 CAACAGAAGGATTCAGGCAATTACCTCTGCCATGCGGTGGAACATGGGTT 1982

 588 rPheArgGluAlaGlnHisTrpGlnLeuProGluAspGlyIleMetA 605
     ::::::        |||        ||| :::||
1983 CATACAAACTCTCTTAAGGTAACCCTG.........GAAGTCATTGACA 2023

 605 laGluHisLeu 608
     :||||||||||
2024 CAGAGCATTG 2034
```

seq_name: gb_sts:G31703

seq_documentation_block:
```
LOCUS       G31703          2530 bp    DNA            STS      28-SEP-1998
DEFINITION  sWSS1973 Eric D. Green Homo sapiens STS genomic, sequence tagged
            site.
ACCESSION   G31703
NID         g1916428
VERSION     G31703.1  GI:1916428
KEYWORDS    STS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2530)
  AUTHORS   Bouffard,G.G., Iyer,L.M., Idol,J.R., Braden,V.V., Cunningham,A.F.,
            Weintraub,L.A., Mohr-Tidwell,R.M., Peluso,D.C., Fulton,R.S.,
            Leckie,M.P. and Green,E.D.
  TITLE     A collection of 1814 human chromosome 7-specific STSs
  JOURNAL   Genome Res. 7 (1), 59-64 (1997)
  MEDLINE   97189344
REFERENCE   2  (bases 1 to 2530)
  AUTHORS   Green,E.D.
  TITLE     Human chromosome 7 STSs (1997)
  JOURNAL   Unpublished (1997)
COMMENT     Synonyms: SEMIII
            GDB_DSEG: SEMIII
            Contact: Eric D. Green
            Genome Technology Branch
            National Human Genome Research Institute/NIH
            49 Convent Dr., MSC4431, Bldg. 49, Rm. 2A08, Bethesda, MD 20892
            Tel: 3014020201
            Fax: 3014024735
            Email: egreen@nhgri.nih.gov
            Primer A: CCCTGGAGTCATTGACACAG
            Primer B: GGTATGTCCTGGCCTTTGCCG
            STS size: 230
            PCR Profile:
                 Presoak:        0 degrees C for  0.00 minute(s)
                 Denaturation:  92 degrees C for  0.17 minute(s)
                 Annealing:     55 degrees C for  1.00 minute(s)
                 Polymerization:72 degrees C for  1.00 minute(s)
                 PCR Cycles:    35
            Thermal Cycler: PerkinElmer 9600
            Protocol:
                 Template:       30-100 ng
                 Primer:         each 1 uM
                 dNTPs:          each 200 uM
                 Taq Polymerase: 0.05 units/ul
                 Total Vol:      10 ul
            Buffer:
                 MgCl2:          1.5 mM
                 KCl:            100 mM
                 Tris-HCl:       10 mM
                 NH4Cl:          5 mM
                 pH:             8.6
```

This STS has been incorporated into the NHGRI chromosome 7
physical map, but was developed by another investigator. See
GenBank record: L26081   For additional information about the NHGRI
chromosome 7 mapping project, see
http://www.nhgri.nih.gov/DIR/GTB/CHR7.   Also see Genomics
11:548-64 (1991) [MUID=92128937].

```
FEATURES             Location/Qualifiers
     source          1..2530
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /map="7"
                     /clone_lib="Eric D. Green"
     gene            1..2530
                     /gene="SEMIII"
     STS             2006..2235
                     /gene="SEMIII"
     primer_bind     2006..2026
                     /gene="SEMIII"
     primer_bind     complement(2215..2235)
                     /gene="SEMIII"
BASE COUNT      786 a      518 c      576 g      650 t
ORIGIN
```

alignment_scores:
```
     Quality:       649.50            Length:    687
     Ratio:         1.620             Gaps:       27
     Percent Similarity: 58.370       Percent Identity: 28.093
```

alignment_block:
US-09-041-236-2 x G31703

Align seg 1/1 to: G31703 from: 1 to: 2530

```
  1 LeuLeuLeuLeuTrpAlaAlaAlaAlaSerAlaAlaGlnGlyHisLeuAr 17
    |||:||:|||: |||:::::::|||:::::::::|||:::||
 34 ATTGTCTGTCTTTCTCGGGAGTATTACTTACGACCAAGGACAAACTATCA 83

 17 gSerGly..........ProArgIlePheAlaValTrpLysGlyHisV 30
    |||:::|||        |||:::|||   :||||
 84 GAATGGGAAGAACAATGTGCCAAGGCTGAAATTATCCTACAAAGAAATGT 133

 30 aIGlyGlnAspArgValAsp..............PheGlyGlnThrGlu 41
    ::|::::|||:::|||                   :::|||||||
134 TGGAATCCAACAATGTGATCACTTCAATGGCTGGCCAACAGCTCCAGT 183

 42 ProHisThrValLeuPheHisGluProGlySerSerValTrpValGl 58
    |||||||||| |||:::|||         ||| :::::::||||||
184 TATCATACCTTCCTTTTGGATGAGGAA...CGGAGTAGGCTGTATGTTGG 230

 58 yGlyArgGlyLysValTyrLeuPheAspPheProGluGlyLysAsnAlaS 75
    |||||||::::  |||||||::    |||||||:::|||
231 AGCAAAGGATCACACATATTTCATTCGACCTGGTTAATATCAAGGATTTC 280

 75 erValArgThrValAsnIleGlySerThrLys................ 85
    ::   |||||||:::|||        :::::
281 AAAAGATTGTGTGGCCAGTATCTTACACCAGAAGAGAGATGAATGCAAGTGG 330

 86 ...GlySerCysLeuAspLysArgAspCysGluAsnTyrIleThrLeuLe 101
    |||:::    :::|||  :||||| :::|||||:::|||
331 GCTGGAAAAGACATCCTGAAA...GAATGTGCTAATTTCATCAAGGTACT 377

101 uGluArgArgSerGlu...GlyLeuLeuAlaCysGlyThrAsnAlaArgH 117
    |:::     |||::::   |||:::|||:::||| :::|||
378 TAAGGCATAATAATCAGATCACTTCACGTACGCCTGTGGAACGGGGCTTTC 427

117 lsProSerCysTrpAsnLeuValAsnGlyThrValProLeuGlyGlu 133
    |||||      ||| :::||:||:::|||
428 ATCCAATTTGC.................ACCTACATTGAAATTGGACAT 459

134 Met............................................ 134
```

```
460 CATCCTGAGGACAATATTTTAAGCTGGAACTCACATTTGAAAACGG      509
135 .ArgGlyTyrAlaProPheSerProAspGluAsnSerLeuValLeuPheG  151
    |||||  ::|||:|||:|||  ::|
510 CCGTGGGAAGAGTCCATATGACCCTAAGCTGCTGACAGCATCCCTTTTAA  559

151 luGlyAspGluValTyrSerThrIleArgLysGlnGluTyrAsnGlyLys  167
    ||||||           :::|||||
560 TAGATGGAGATTATACTCTGGAACTGCAGCTGATTTTATGGGGCGAGAC   609

168 IleProArgPheArgArgIleArgGlyGluSerGluLeuTyrThr.....  182
    ::|       ||||    :::    |||
610 TTTGCTATCTCCGAACTCTGGGACCACCACCCAATCACAGGACGAGCA    659

183 ....SerAspThrValMetGlnAsnProGlnPheIleLysAlaThrIleV  198
    :::                 |||||:|||||
660 GCATGAATTCCAGGTGGCTCCAATGATCCAAAGTTCATTAGTGCCACCTCA 709

198 alHisGln...AspGlnAlaTyrArgAspLysIleTyrTyrPhePheArg  213
    ::  :::|    |||       ::::::|:|||:|||||
710 TCTCAGAGAGTGACAATCCTGAAGATGACAAAGATACTTTTCTTCCGT    759

214 GluAspAsnProAspLysAsnGlyGluAlaProLeuAsnValSerArgVa  230
    |||:::|:::|||:    :||  ::|||:|||||
760 GAAAATGCAATAGATGGAGAACACTCTGGAAAAGCTACTCACGCTAGAAT  809

230 lAlaGlnLeuCysArgGlyAspAspGlyGlySerSerLeuSerValS     247
    ||||:  :|||||:||||||:|||||||||||
810 AGGTCAGATATGCCAAGAATGACTTTGGAGGGCACAGAGTCTG...GTGA  856

247 erLysTrpAsnThrPheLeuLysAlaMetLeuValCysSerAspAlaAla  263
    ::||||||:|||||||||||   |||  |||||::::|||
857 ATAAAATGGACAACATTCCTCAAAGCTGCGTCTGATTTGCTCAGTGCCAGGT 906

264 ThrAsn.......LysAsnPheAsnArgLeuGlnAspValPheLeuLe    277
    |||         ::::|||||||||||||||||||
907 CCAAATGGCATTGACACTTCATTTGAATGACTGCGAGATGTATTCCTAAT  956

277 u......ProAspProSerGlyGlnTrpProArgAspThrArgValTyrGlyV 292
    |||          |||||||||||||||
957 GAACTTTAAAGATCCT........AAAAATCCAGTGTATATGGAG        994

292 alPhe........SerAsnProTrpAsnTyrSerAlaValCysValTyr   305
    ||||         :|||||:|||||:::|||
995 TGTTTACGACTTCCAGTAACATTTTCAAGGGATCAGCCGTGTGTATGTAT  1044

306 SerLeuGlyAspIleAspLysValPhe..............           314
    ::|||:|||:::  ::|||||
1045 AGCATGAGTGATGTGAGAAGGGGTGTTCCTTGGTCCATATGCCCCACAGGGA 1094

315 ....ArgThrSerLeuLysGlyTyrIleSerSerLeuProAsnProA     330
    :::    |||||||||||    :::|||||
1095 TGGACCCAACTATCAATGGGGTGCCCTATCAAGGAGAGTCCCTATCCAC   1144

330 rgProGlyLysCys............LeuProAspGlnGln            339
    ||||||||:  |||             :::
1145 GGCCAGGAACTTGTCCCAGCAAAACATTGGTGTTTGACTCTACAAAG     1194

340 ProIleProThrGluThrPheGlnValAlaAlaAspArgHisProGluValAl 356
    :::|||||                      |||||
1195 GACCTTCCTGATGATGTTATAACCTTGCAAGAAGTCATCCAGCCATGTA   1244

356 aGlnArgValGluProMetGly.........ProLeuLysThrProL     369
    ||||                 :::     ::|||||
1245 CAATCCAGTGTTCCTATGAACAATCGCCCAATAGTGATCAAAACGGATG   1294

369 euPheHisSerLysTyrHisTyrGlnLysValAlaValHisArgMetGln  385
    ::   :|||||:|  ::||||||||||:|
1295 TA.......AATTATCAATTACACAAATTGTCGTAGACCGAGTGGAT     1335

386 AlaSerHisGlyGluThrPheHisValLeuTyrLeuThrAspArgArgGl  402
    ||:::|||          |||||:|||:||||||||   ||||
1336 GCAGAAGATGGACAG..TATGATGTTATGTTTATCGGAACAGATGTTGG   1382
```

```
402 yThrIleHisLysValValGlu...ProGlyGluGlnGluHisSerPheA  418
    |||:||||::|||||||||||   |||   |||  ::::::||||||:
1383 GACCGGTTCTTAAAGTAGTTCAATTCCTAAGGAGACTGGTATGATTTAG   1432

418 laPheAsnIleMet...GluIleGlnProPheArgArgAlaAlaAlaIle  433
    |||::|||:||||    |||:|||::|||||||||:|||:|
1433 AAGAGGTTCTGCTGGAAGAAATGACAGTTTTCGGGAACCGACTGCTATT   1482

434 GlnThrMetSerLeuAspAlaGluAlaArgLysLeuTyrValSerSerGl  450
    :::|||:||||:::|||||::||||||||||||:::|||:::||||||
1483 TCAGCAATGGAGCTTTCCACTAAGCAGCAACAACTATATATGGTTCAAC   1532

450 nTrpGluValSerGlnValProLeuAspLeuCysGluValTyrGlyGlyG  467
    ::|||:::||||:||||:::|||||:::||||||:::|||||:||||||
1533 GGCTGGGGTTGCCCAGCTGCCCCTTACACCGGTGTGATATTTACGGGAAAG 1582

467 lyCysHisGlyCysLeuMetSerArgAspProTyrCysGlyTrpAspGln  483
    :::|||:::||||||:::||||||:::|||:::|||:::||||||||||
1583 CGTGTGCTGAGTGTTGCCTCGCCGGACCCCTTACTGTGCTTGGGATGGT   1632

484 GlyArgCysIleSerIleTyr.....SerSerGluArgSerValLeuGl   498
    :::|||:::|||:::||||||     :::||||||:::|||||||:||
1633 TCTGCATGTTCTCGCTATTTCCCACTGCAAAGAGACGCACAAGACGACA   1682

498 nSerIleAsnProAlaGluProHisLysGluCysProAsnProLysProA  515
    ::|||:::|||:::|||||:::|||||:::|||:::|||:::|||:||
1683 AGATATAAGAAATGGAGACCCACTGACTCAGTTCAGACTTACACCATG    1732

515 spLys...........AlaProLeuGlnLeuValSerValSerLeu...AlaPro 526
    :||:           ||||||:::|||:::||||||:::|||    ||||||
1733 ATAATCACCATGGCCACAGCCCTGAAGAGAGAATCATCTATGTGTAGAG   1782

527 AsnSerArgTyrTyrLeuSerCysProMetGluSerArgHisAlaThrTy  543
    |||:::||||||||||:::|||:::|||||||:::|||||:::|||||
1783 AATAGTAGCACATTTTGGAATGCAGTCCGAGTCGCAGAGAGCGCTGGT    1832

543 rSerTrpArgHisLysGluAsnValGluGlnSerCysGluProGlyHisG  560
    :::|||||||||||:::|||||:::|||:::||||||:::|||||||:
1833 CTATTGGCAATTCCAGAGGCGAAATGAAGAGCGAAAAGAAGAGATCAGAG  1882

560 lnSerProAsnCysIle...............LeuPheIleGluAsnLeu  571
    :::|||:::|||:::|||               |||||||||::|||
1883 TGGATGATCATATCATCAGGACAGATCAAGGCCTTCTGCTACGTAGTCTA  1932

572 ThrAlaGlnIleTyrGlyHisTyrPheCysGluAlaGlnGlyGlySerTy  588
    |||:::|||:::||||||:::|||:::|||||||:::|||||:::|||
1933 CAACAGAAGGATTCAGGCAATTACCCTGCCATGCGGTGGAACATGGGTT   1982

588 rPheArgGluAlaGlnIleHisTrpGlnLeuProGluAspGlyIleMetA  605
    :::|||:::|||||:::|||:::|||||||:::        :::|||||
1983 CATACAAACTCTTCTTAAGGTAACCCTG........GAAGTCATTGACA   2023

605 laGluHisLeu                                         608
    :::||||||
2024 CAGAGCATTTG                                        2034
```

seq_name: gb_ro:MUSC1

```
JOURNAL    Submitted (02-MAY-1996) to the DDBJ/EMBL/GenBank databases.
           Masahiko Taniguchi, Institute for Developmental Research, Aichi
           Prefectural Colony, Department of Morphology; Kamiya-cho Kasugai,
           Aichi 480-03 Japan, Kasugai, Aichi 480-03, Japan
           (E-mail:taniguti@nips.ac.jp, Tel:0568-88-0811, Fax:0568-88-0829)
REFERENCE  2 (bases 1 to 2356)
AUTHORS    Taniguchi,M.
TITLE      Mouse collapsin-1
JOURNAL    Unpublished (1996)
FEATURES             Location/Qualifiers
     source          1..2356
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
     CDS             38..2356
                     /codon_start=1
                     /product="collapsin-1"
                     /protein_id="BAA19773.1"
                     /db_xref="PID:d1020054"
                     /db_xref="PID:g2055315"
                     /db_xref="GI:2055316"
                     /translation="MGWFTGIACLFWGVLLTARANYANGKNNVPRLKLSVKEMLESNN
                     VITFNGLANSSSYHFLLDEERSRLYVGAKDHIFSFNLVNIKDFQKIVWPVSYTRRDE
                     CKWAGKDILKECANFIKVLEAYNQTHLYACGTGAFHPICTYIEVGHHPEDNIFKLQDS
                     HFENGRGKSPYDPKLLTASLLIDGELYSGTAADFMGRDFAIFRTLGHHPIRTEQHDS
                     RWLNDPRFISAHLIPESDNPEDDKVFFFRENAIDGEHSGRATHARIGQICKNDFGGH
                     RSLVNKWTTFLKARLICSVPGPNGIDTHFDELQDVFLMNSKDPKNPIVYGVFTTSSNI
                     FKGSAVCMYSMSDVRRVFLGPYAHRDGPNYQWVPYQGRVPYPRPGTCPSKTFGGFDST
                     KDLPDDVITFARSHPAMYNPVFPINNRPIMIKTDVNYQFTQIVVDRVDAEDGQYDVMF
                     IGTDVGTVLKVVSVPRETWHDLEEVLEEMVFREPTTISAMELSTKQQQLYIGSTAG
                     VAQLPLHRCDIYGKACAECCLARDPYCAWDGSSCSRYFPTAKRRTRRQDIRNGDPLTH
                     CSDLQHHDNHHGPSLEERIIYGVENSSTFLECSPKSQRALYWQFQRRNEDRKEEIRM
                     GDHIIRTEQGLLLRSLSQKKDSGNYLCHAVEHGFMQTLLKVTLEVIDTEHLEELLHKDD
                     DGDGSKIKEMSSSMTPSQKVWYRDFMQLINHPNLNTMDEFCEQVWKRDRKQRQRPGH
                     SQGSSNKWKHMQESKKGRNRRTHEFERAPSV"
BASE COUNT      708 a    525 c    563 g    560 t
ORIGIN

alignment_scores:
       Quality:      635.00          Length:        688
       Ratio:          1.591         Gaps:           27
       Percent Similarity: 57.994    Percent Identity: 27.907

alignment_block:
US-09-041-236-2 x MUSC1

Align seg 1/1  to: MUSC1  from: 1  to: 2356


    4  LeuLeuTrpAlaAlaAlaAlaSerAlaGlnGlyHisLeuArgSerGly..   19
       |||:||:|| :::::|:   :|||:  |||
   65  CTTTTCTGGGGTGTATTACTTACCAGCCAGAGCAAACATGCAAACGGAAA  114

   20  .........ProArgIlePheAlaValTrpLysGlyHisValGlyGlnA   33
       ||||||:::::: : .:::|||
  115  GAACAATGTGCCCAAGACTGAAATTATCCTACAAAGAAATGTTGGAATCCA  164

   33  spArgValAsp............PheGlyGlnThrGlyProHisThr     44
       :::|||:|||           |||||
  165  ACAATGTGATCACTTTTAATGGCTTGGCCAACAGCTCCAGTTACCACACC  214

   45  ValLeuPheHisGluProGlySerSerValTrpLysValGlyArgGl     61
       |||:||:: ::|||| ::::::|::::|
  215  TTCCCTTCTGGATGAAGAA...CGGAGTAGACTATATGTTGGAGCAAAAGA  261

   61  yLysValTyrLeuPheAspPheGluGlyLysAsnAlaSerValArgT     78
       :::|||||: ||| ::: :::|:
  262  TCATATATTTTCATTCAACTTGGTGAACATTAAAGANTTTCAAAAGATTG  311

   78  hrValAsnIleGlySerThrLys............GlySer          87
       |:::::|||           ||:||
  312  TGTGGCCAGTATCTTACACAAGGAGAGATGAATGCAAATGGGCTGGAAA  361

   88  CysLeuAspLysArgAspCysGluAsnTyrIleThrLeuLeuArgAr    104
       ::: :::|||  || |:: :: :: :||
  362  GATATCCTGAAA...GAATGTGCCAATTTCATCAAGGTCTGGAGGCTTA  408

  104  gSerGlu..GlyLeuLeuAlaCysGlyThrArgAlaArgHisProSerC   120
       :||||   :: |||
  409  TAATCAGACTCACTTGTATGCCTGTGGACTGGGCTTTCCATCCAATCT  458

  120  ysTrpAsnLysTrpThrThrPheLeuLysAlaArgLeu.........   134
       ||: ::: ::::::
  459  GC............ACCTATATTGAAGTTGGACATCATCCTGAG  490

  135  ..............ArgGlyTy                            137
                        |||||
  491  GACACATTTTAAGCTGCAGGACTCACATTTTGAAACGGTCGTGGGAA  540

  137  rAlaProPheSerProAspGluAsnSerLeuValLeuPheGluGlyAspG  154
       ::: ||| :::
  541  GAGCCCTTATGATCCCAAACTACTGACTGCCTCCTTCTAATAGACGGTG  590

  154  luValTyrSerThrIleArgLysGlnGluTyrAsnGlyLysIleProArg  170
       |::::|||||                 :::
  591  AGTTGTACTCTGGAACTGCTGCGGACTTCATGGGACGGGACTTCGCTATC  640

  171  PheArgArgIleArgGlyGluSerGluLeuTyrThr.........SerAs  184
       ||||| :::              ::: :::
  641  TTCAGAACACTGGGGCACCATCACCCCATCAGGAGCGGAGCAGCATGACTC  690

  184  pThrValMetGlnAsnProGlnPheIleLysAlaThrIleVal...HisG  200
       :|:|:|||:::
  691  CCGGTGGCTCAATGATCCTAGATTCATCAGTGCCCATCTCATCCCAGAGA  740

  200  lnAspGlnAlaTyrAspAspLysIleTyrPheArgGluAspAsn       216
       ::||::   |||||||||||::||||||||::: :::
  741  GTGACAACCCTGAAGATGACAAAGTATATTTTTCTTCCGAGAAAATGCA  790

  217  ProAspLysAsnProGluAlaProLeuAsnValSerArgValAlaGlnLe  233
       ||| :: ||||:::  ::|||:::::::::::
  791  ATAGACGGAGAACACTCTGGAAAAGCCACTCATGCTAGAATAGGTCAGAT  840

  233  uCysArgGlyAspGlnIleGlyLeuSerSerLeuSerValSerTrpTrpA  250
       :|:::|| ::::: ::::|:: |||
  841  ATGCAAGAGATGACTTTGGTGGACACAGAAGTCTT...GTGAATAAATGGA  887

  250  snThrPheLeuLysAlaMetLeuValCysSerAspAlaAlaThrArgAsn. 265
       :::|||||||:|:::: ||| |||
  868  CAACATTCCTAAAAGCACACGCCTGATTTGCTCTGTGCCCGTCCCAATGGC  937

  266  ......LysAsnPheAsnArgLeuGlnAspValPheLeuGlu........Pr 278
       |||:|||:::::::|||||||:::
  938  ATTGACACCCATTTTGATGAATTGCAGGATGTATTCCTATGAACTCTAA  987

  278  oAspProSerGlyGlnTrpArgAspThrArgValTyrGlyValPhe....  293
       :|::: ||| |||||||:::::
  988  AGATCCT.........AAAAATCCGATCGTCTATGGAGTGTTCACAA  1025

  294  .....SerAsnProTrpAsnTyrSerAlaValCysSerValTyrSerLeuGly 308
       :::|||||| |||:::||::|||||:::::|:::
 1026  CATCAGGAGCAACATCTTTAAGGGATCTGCGTGTGCATGTACAGCATGAGT  1075

  309  AspIleAspLysValPhe........................ArgTh   316
       :::|||||||:::
 1076  GATGTAAGAAGGGTGTTCCTTGGTCCATATGCTCACAGAGATGGTCCCAA  1125

  316  rSerSerLeuLysGlyTyrHisSerSerLeuProAsnProArgProGlyL  333
       |||||||||||||||||||||:|||
 1126  CTATCAGTGGGTGCCTTACCAAGGAAGAGTCCCTTATCCACGGCCAGGAA  1175

  333  ysCys.........................LeuProAspGlnGlnProIlePro 342
       |||                          ::||||
 1176  CTTGTCCCAGTAAAAACATTTGGCGGATTGACTCCAGAAACGGACCTTCCT  1225

  343  ThrGluThrPheGlnValAlaAspArgHisProGluValAlaGlnArgVa  359
       :::::::|||  ||||||||  ::: :||  ||
```

```
LOCUS        MMRNASEMA        2872 bp     mRNA              ROD      08-JUL-1996
DEFINITION   M.musculus mRNA for semaphorin A.
ACCESSION    X85990
NID          g854323
VERSION      X85990.1  GI:854323
KEYWORDS     semA gene; semaphorin.
SOURCE       house mouse.
  ORGANISM   Mus musculus
             Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
             Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE    1  (bases 1 to 2872)
  AUTHORS    Puschel,A.W., Adams,R.H. and Betz,H.
  TITLE      Murine semaphorin D/collapsin is a member of a diverse gene family
             and creates domains inhibitory for axonal extension
  JOURNAL    Neuron 14 (5), 941-948 (1995)
  MEDLINE    95267431
REFERENCE    2  (bases 1 to 2872)
  AUTHORS    Adams,R.H.
  TITLE      Direct Submission
  JOURNAL    Submitted (30-MAR-1995) R.H. Adams, Max-Planck-Institute fuer
             Hirnforschung, Deutschordenstr. 46, D- 60528 Frankfurt, FRG
COMMENT      Sequence overlapping with those under the acc#U02528, L26080,
             L26082.
FEATURES             Location/Qualifiers
     source          1..2872
                     /organism="Mus musculus"
                     /strain="NMRI"
                     /db_xref="taxon:10090"
                     /dev_stage="12 days embryo"
                     /clone_lib="Stratagene lambda ZAP"
                     /clone="semag21"
                     /clone="semac3"
     gene            210..2456
                     /gene="semA"
     CDS             210..2456
                     /gene="semA"
                     /codon_start=1
                     /product="semaphorin A"
                     /protein_id="CAA59982.1"
                     /db_xref="PID:g854324"
                     /db_xref="GI:854324"
                     /db_xref="MGD:MGI:107561"
                     /db_xref="SPTREMBL:Q62177"
                     /translation="MGRAEAAVMIPGLALLWVAGLGDTAPNLPRLRLSFQELQARHGV
                     RTFRLERTCCYEALLVDEERGRLFVGAENHVASLSLDNISKRAKLAWPAPVEWREEC
                     NWAGKDIGTECMNFVRLLHAYNHTHLLACRTGAFHPTCALWRWATAGGTHASTGPEKL
                     EDGKGTPYDPRHRPPSVLVGEELYSGVTADLMGRDFTIFRSLGQNPSLRTEPHDSRW
                     LNEPKFVKVFWIPESENPDDKIYFFRESAVEAAPAMGRMSVSRVGQICRNDLGGQR
                     SLVNKWTTFLKARLVCSVPGVEGDTHFDQLQDVFLLSSRDRQTPLLYAVFSTSSGVFQ
                     GSAVCVYSMNDVRRAFLGPLPHKEGPTHQWVSYQGRVPYPRPGMCPSKTFGTFSSTKD
                     FPDDVIQFGRNHPLMYNPVLPMGGRPLFLQVGAGVTFTQIAADRVAAADGHYDVLFIG
                     TDVGTVLKVLSVPKGRRPNSEGLLLEELQVFEDSAAITSMOISSKRQQLYASRAVA
                     QIALHRCTALGRACAECCLARDPYCAWDGSACTRFQPTAKRFRRODIRNGDPSTLCS
                     GDSSHSVLLEKKVLGVESSSAFLECEPRSLQAHVOWTFQGAGEAAHTQVLAEERVERT
                     ARGLLLRGLRRQDSGVYLCVAVEQGFSQPLRRLVLHVLSAAQAERLARAEEAAAPAPP
                     GPKLWYRDFLQLVEPGGGGGANSLRMCRPQPGHHSVAADSRKGRNRMHVSELRAER
                     GPRSAAHW"
BASE COUNT     592 a    799 c    882 g    599 t
ORIGIN

alignment_scores:
     Quality:   634.50                    Length:     747
     Ratio:       1.571                    Gaps:        27
Percent Similarity:  54.083    Percent Identity:  27.443

alignment_block:
US-09-041-236-2 x MMRNASEMA

Align seg 1/1  to: MMRNASEMA     from:  1  to: 2872
```

```
1226  GATGATGTCATAACTTTGCGAAGAGTCATCCAGCCATGTCAACCCAGT  1275
 359  1GluProMetGly...........ProLeuLysThrProLeuPheHisS
       |   |:|::::::             |:||||:|| |||
1276  GTTTCCTATAAATAATCGCCCGATCATGATCAAAACAGATGTA.......  1318
 372  erLysTyrHisTyrGlnLysValAlaValHisArgMetGlnAlaSerHis
       ::||||:|::|||       :|:||  |||:||:::::
1319  ..AATTATCAGTTCACACAAATTGTGTAGACCGAGTGGATGCAGAAGAT  1366
 389  GlyGluThrPheHisValLeuTyrLeuThrThrAspArgGlyThrIleH1  405
       |||:||:   :: |||  ||:|:|:|||:||  ||||||  ||||||::
1367  GGCCAG...TATGATGTTATGTCACCGGAACAGATGTTGGAACCCTTCT  1413
 405  sLysValGlu...ProGlyGluGlnHisSerPheAlaPheAsnI  421
       ||||| ::|:    |:::| |||:|:|||||:|:|:|  |
1414  TAAAGTGGTTTCAGTCCCCAAGGAGACTTGGCATGACCTAGAGAAGTTC  1463
 421  leMet...GluIleGlnProPheArgArgAlaAlaAlaIleGlnThrMet  436
       :::::|  ||| ||||:|:|||| ||||||:|:|::|||:|:|::||:|||
1464  TTCTGGAAGAATGACCGGTCTTCCGGGAACCAACACTATTCGGCAATG  1513
 437  SerLeuAspAlaGluArgArgLysLeuTyrValSerSerGlnTrpGluVa  453
       :::|:|||||||:|:||| ||:||:|||:||||||||||||||||||||
1514  GAGCTTCTACTAAACAGCACAGCTGTACATTGGCCTCAACTGCCGGAGT  1563
 453  lSerGlnValProLeuAspLeuCysGluValTyrGlyValGlyGlyCysHissG  470
       | :|:|||:|||||||||||:|||:|:|||||||:|:|||||:|:|:|||
1564  GGCACAGCGTTCCCTCTACACCCGTCTGACACATCTGCAAAGCCTGTCAG  1613
 470  lyCysLeuMetSerArgAspProTyrCysGlyTrpAspGlnGlyArgArgCys  486
       || ::|::::|:||||::|:||||:||||:||||:::||
1614  AATGCTGCCTCGCTCGGGACCCTTACTGTCGCCTGGGATGGGTCCTCATGC  1663
 487  IleSerIleTyr......SerSerGluArgSerValLeuGlnSerIleLeuAs  501
        ::||   ||      |:::|||||:||||||||||||||||:| :||
1664  TCACGCTATTTCCTACTGCAAAGAGGCCCACAAGACGACAAGATATAAG  1713
 501  nProAlaGluProHisLysGluCysProAsn..................  511
       : :::|:||||      :|::|||||
1714  GAATGGAGAGACCCACTGACTCTCTGACTTGCAGCACCATGATAATC  1763
 512  ........ProLysProAspLysAlaProLeuGlnLysValSerLeuValAla  525
               :| :::::       |:::::   |:::|:| |||
1764  ATCATGGGCCCAGCCTTGAAGAGAGAATCATCTATGGAGTGGAA......  1807
 526  ProAsnSerArgTyrTyrLeuSerCysProMetGluSerArgHisAlaTh  542
       |||||           ::|||  || ::||:|| |||:::||||:::|||
1808  ..AACAGTAGTACGTTCTGGAATGCAGTCCGAAGTCACAGAGAGCCCTT  1854
 542  rTyrSerTrpArgHisLysGluAsnValGlnSerCysGluCysProGlyH  559
       ||||          |:::|:|||      |:|||:|:|||||:|:|:|
1855  GGTATATTGGCAAATTCAGAGGAGAAAATGAAGATCGAAAAGAGGAGATCA  1904
 559  isTyrSerTrpArgHisLysGluAsnValGlnSerCysGluCysProGlyH  559
       :::|||                  |||:|||:|:
1905  GAATGGGTGATCATCATCAGGACAGAACAAGGGCTCCTGCTCCGTAGC  1954
 570  isGlnSerProAsnCysIle..............LeuPheIleGluAsn  570
       :: ||::|  |||:||:::||||:|  |||  ||
1955  CTGCCAGAAGAAGGATTCAGGCAATTACCCTGTGTCACGCCTGGAACACGG  2004
 587  rTyrPheArgGluAlaGlnHisTrpGlnLeuProGluAspGlyIleM  604
        |||:|             :::||| ||   ::: |||
2005  ATTCATGCAAACTCTTCTTAAGGTAACCCTG........GAAGTCATTG  2045
 604  etAlaGluHisLeu  608
       :|||||||||||
2046  ACACAGAACATTG  2059
```

seq_name: gb_ro:MMRNASEMA

seq_documentation_block:

Wed Jan 19 11:15:38 2000

us-09-041-236-2.LEE.rge

```
 30 TGGGGCGGCTGCTGCCAATGCAGGGTGACCCTGAGGGTGTGCCCGGTGT  79
 19 .GlyProArgIlePheAlaValTrpLysGlyHisValGlyGlnAspArgV  35
        :::::::::   ::::|||  ::::||      :::::|||
 80 GGGTGTGCTCTGATTGTGGCCAGGCGGGCACCCTTGGTGGAGAGGAAA   129

 35 aIAspPheGlyThr.......... .......GluProHisThr...     44
        ::   ||||||||           :::::|||||||
130 CGATCCTGGGGCCAGAGTCCAGGACGTCCTGAAGGCTCCTCCACACGCCT 179

 44 ...............................AspAlaAlaThrAsnLysAsnPheAs  44
                                     ||| ::  ::  ::::|||
180 GCTGAACCCTGACGCGCTGAGCTGCCGCCGCATGGGGCGGGCTGAGGCGC  229

 45 .ValLeuPheHisGluProGlySerSerValTrpValGlyGlyArgG    61
        :::::::   ::   :::::|||||::::|||
230 CGTCATGATC......CCAGGGCTGGCCCTTCTCGGGGTAGCAGGGCTAG  273

 61 ..ly.............................................  61
    ::|||
274 GGGATACTGCCCTAACCTTCCCCGCCTTCGGCTCTCCTTTCAAGAATTA  323

 61 ...............................................  61
324 CAGGCCCGACCATGGTGTCCGAACCTTCAGCTGGACGCGACCTGCTGTTA 373

 62 .LysValTyrLeuPheAspPheGluPheGlyLysAsnAlaSerValArgT   78
    ::::::||| |||   :::|||||||||||||:::|||
374 TGAAGCCCTTGCTCTGGGATGAGGACGTGGACCCTGTTGTGGGTGCTG   423

 78 hrValAsnIleGlySerThrLysGlySerCysLeuAspArg.......   92
    ::::::|||   ::  ||::||||||::::|||
424 AGAACCACGTGGCTTCCCTCAGCCTGGACAACATCAGCAAGCGAGCCAAG  473

 92 ...............................................  92
474 AAGCTGGCCTGCGCCCGCCCCGGGCTGGAAATGGCGTGAAGAATGCAACTGGGC  523

 93 ...........................AspCysGluAsnTyrIleThrLeuLeuGluA  103
                              :::|||  |||::||   ||||||::::|||
524 AGGGAAGGACATTGGTACCGAGTCCGAGCGTCATGTACACACCAGTCCTGCCTGCACG  573

103 rgArgSerGlu.. ..GlyLeuLeuAlaCysGlyThrAsnAlaArgHisPro   118
    ::||  |||     ||||||||::  |||   |||  ::::|||
574 CCTACAACACCACCCACTTGCTGCCGTGCTGGACAGGGGCTTTCCACCCA   623

119 SerCys.........TrpAsnLeuValAsnGlyThrValProLe       131
    ::::|||        |||  |||||::||||||::
624 ACCTGCGCCTGTGAGGTGGCCACCCTGGAGGAACCCATGGTTCAAC   673

131 uGly................GluMetArgGlyTyrAlaProPheSerProA  143
    :::||               :::|||||||   ::::|||||::::|||
674 TGGACCGGAGAAACTTGAGACGCCAAGGGAAGACTCCTTATGACCCAA   723

143 spGluAsnSerLeuValLeuPheGluGluValAspGluValTyrSerThrIle  159
    ::::||| :::|||||::||||||::  |||::::||||||||||:::
724 GGCATCGGCCTCCCCTGCTGTGCGGGGGAAGAAACTGTATTCTGGGGTG  773

160 ArgLysGlnGluTyrAsnGlyLysIleProArgPheArgArgIleArgGl  176
    :::|||  ::::::|||||||||::::|||
774 ACAGCAGACTTATGGGCGGCGGGACTTTACCAATCTTTCGAAGCCTTGGTCA  823

176 yGluSerGluLeuTyrThr...........SerAspThrValMetGlnAsnP  190
    :::|||   :::||     ::::::::|||::::||
824 GAATCCGAGTCTCCGAACAGAGCCCCATGATTCCCGCTGGCTCAATGAAC  873

190 roGlnPheIleLysAla...ThrIleValHisGlnAspGlnAlaTyrAsp  205
    ||||||||||||||||     :::||||||::::||::::|||:::|||
874 CCAAGTTTGTCAAGGTCTTTGGATCCCAGAGAGTGAGAACCCTGATGAC  923

206 AspLysTyrTyrPhePheGluArgGluAspAsnProAspLysAsnProGl  222
    ::::::||||||||||::::::::::::||||||:::::::||
924 GATAAAATCTATTTCTTCTCCCGAGTCCGCTGTGGAGCAGCACCAGC   973
```

```
222 u...AlaProLeuAsnValSerArgValAlaGlnLeuCysArgGlyAspG  238
        :::||  ::  ::::::|||::::::::|||::::|||::::||
974 AATGGCAGCCATGTCTGTCGTGTTGGCCAGATCTGCAGGAATGACC  1023

238 lnGlyGluSerSerLeuSerValSerLysTrpAsnThrPheLeuLys    254
    ||||||::  :::||::::::||||||::: ::::::|||||||||
1024 TGGGTGGCCAGCGGCGAGCTTG...GTCAACAAATGACCACCATTTCTGAAG  1070

255 AlaMetLeuValCysSer.......AspAlaAlaThrAsnLysAsnPheAs  269
    ||||:::||::::::     :::::::::||   ::  :::::
1071 GCGGGCGGCTTGTGTGTCAGTACCTGGAGTTGAGGGTGACACCCACTTTGA  1120

269 nArgLeuGlnAsnAspValPheLeuLeuProAspProSerGlyGlnTrpArgA  286
    :::::||||||||||::::::::::|||::::|||  ::::
1121 CCAACTTCAGGATGTTTCCTTCTG........TCCTCCCGAGACCCGCC  1161

286 spThr...ArgValTyrGlyValPheSerAsnPro..........TrpAsn  298
        ::|||  :::|||::::|||::::::::::||:::
1162 AGACACCTCTCTCTATGCTGTCTTCTCCACCTCCAGTGTGTCTCCAG   1211

299 TyrSerAlaValCysValTyrSerLeuGlyAspIleAspLysValPhe.   314
    ||||||||||||||||||||::::||||||::::::|||::::
1212 GGCTCTGCTGTGTGCGTGTACAGCCTGGGCGACATTGACAAGGTGTCAGCCCTTCTT  1261

315 ......................ArgThrSerLeuLysGlyTyrH      323
                          :::|||||||||::::::
1262 GGGACCTTTGCCTCACAAAGAGGGGCTACACACCAGTGGTGTCCTACC  1311

323 isSerSerLeuProAsnProArgProGlyLysCys.............   334
    ||::::::|||||||::  |||||||::::|||
1312 AGGGTCGTGTCCCCTACCCAAGACCTGCCGCCATGTGCCCAGCAAGACCTTT  1361

335 .......LeuProAspGlnGlnProIleProThrGluThrPheGlnIleValAl  349
           ::|||||||||||::::||||||::::|||
1362 GGGACCTTCAGCTCCCACCAAGGACTTCCCAGATGACGTTATCCAGTTTGG  1411

349 aAspArgHisProGluValAlaGlnArgValGluProMetGlyProLeuL  366
    ::||::::::::|||||||||::::|||::::||||||
1412 TCGGAACCACCCTCTCATGTACACACCAGTCCTGCCCATGGGG.......G  1455

366 ysThrProLeuPhe..........HisSerLysTyrHisTyrGlnLysVal  379
    ||||||||||||            ::::::||||||||||||||
1456 GGGCCCCTCTTCTTCCTACAGTGGAGCTGGGAGTACACCTTCACCCAAATC  1505

380 AlaHisArgMetGlnAlaSerHisGlyGluThrPheHisValLeuTy    396
    ||::  ::  :::|||::::||::::||||||
1506 GCCCAGACCGAGTAGCAGCGCCGATGGACAC...TACGATGTTCTCTT  1552

396 rLeuThrPheAspArgGlyThrHisLysVal.....ValGluProG     411
    :::||||||::::::::|||    ||||||
1553 CATTGGTACAGATGGGCACAGTGGCTGAAAGTGATCCAGTCCCCAAAG  1602

411 lyGluGlnGluHisSerPheAlaPheAsnIleMetGlyIleGlnProPhe  427
    ||::::::||::::||||||::::|||::::||||||
1603 GCAGGCGACCTAATTCTGGAAGAGCTGCAGGTGTTC  1652

428 ArgArgAlaAlaAlaIleAlaIleGlnThrMetSerLeuAspAlaGluArgLy  444
    :::::::||||||||::  :::|||::::::|||::::
1653 GAGGACTCTGCCCTATCACCAGCATGCAAATCCTCTAAAAGGCAACA  1702

444 sLeuTyrValSerSerGlnTrpGluValSerGlnValProLeuAspLeuC  461
    :::||||||||||::  :::||||||::::|||::::::
1703 ACTCTACGGTAGCATCGCGCAGCAGTGGCCAGATTGCTTTGCATCGCT  1752

461 ysGluValTyrGlyGlyCysHisGlyCysLeuMetSerArgAspPro    477
    ||::::::||||||||||::  ||||||::::||::::
1753 GCACGTGCCCTAGGCCGCCGCAGAATGCTGTTGGCCGGTGATCCT  1802

478 TyrCysGlyTrpAspGlyGlnGlyArgCysIleSerIleTyr.....SerSe  492
    ::::::::::|||||||||||::::::||||||::::::
1803 TACTGCGGCTGGGATGGATCAGCTTGCACACGCTTGCACACGGCCAA  1852
```

```
 492 rGluArgSerValLeuGlnSerIleAsnProAlaGluProHisLysGluC  509
     ::::|||      |||::||:|::    ::::||||         |
1853 GAGACGGTTCCGGAGGCAAGACATAAGGAATGGCGACCCCAGCACCCTAT 1902

 509 ysProAsnProLysProAspLysAlaProLeuGlnLysValSerLeuAla  525
     ||  ::   :::   |||::|||  |||:||   |||::::
1903 GCTCTGGAGACTCTTCTCACTCTGCTGCTGGTGCTGGAGAAGAAGGTTGGGT 1952

 526 ProAsnSer...ArgTyrTyrLeuSerCysProMetGluSerArgHisAl  541
     :::|||        :::|||      :::||| |||::|
1953 GTGGAGAGCCGGCAGCGCGTTTCTGGAGTGTGAGCCCGGCTCGCTCCAGGC 2002

 541 aThrTyrSerTrpArgHis................................  547
     |  :::|||
2003 GCATGTGCAGTGGACCTTCCAAGGTGCAGGGGAGGCAGTCACACCCAGG 2052

 548 .......LysGluAsnValGluGlnSerCysGluProGlyHisGlnSer  561
     :::||||:::||||||:::::::::
2053 TGCTGGCTGAGGAGAGTAGAGCGCACTGCGCGGGGG.............. 2090

 562 ProAsnCysIleLeuPheIleGluAsnLeuThrAlaGlnGlnTyrGlyHi  578
     ||||::      |||       ||
2091 ..........CTGCTGTTGCGGGGGCCTCGCGCCCAGGATTCTGGGCGT 2128

 578 sTyrPheCysGluAlaGlnGluGlyGlySerTyrPheArgGluAlaGlnHisT  595
     |||::::|    |||   |||     :::::::
2129 GTATCTAATGCGTCGCGGTTGAACAAGGCTTTTCACAACCACTGCGTCGCC 2178

 595 rpGlnLeuLeuProGluAspGlyIleMetAlaGluHisLeu  608
     |||       :::||||||:::|||
2179 TGGTGCTGCATGTGTTGAGTGCGGCGCAGGCTGAACGACTG 2219
```

OM of: US-09-041-236-2 to: N_Geneseq_36:*   out_format : pfs

Date: Jan 18, 2000 9:18 PM

About: Results were produced by the GenCore software, version 4.5,
    Copyright (c) 1993-1998 Compugen Ltd.

Command line parameters:
```
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09041236/runat_15012000_164052_532/app_query.fasta.1
-DB=N_Geneseq_36 -QFMT=fastap -SUFFIX=LEE.rng -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
-MAXLEN=1000000 -USER=US09041236 -NCPU=3 -ICPU=3 -NO_XLPXY -WAIT
-THREADS=1
```

Search information block:
Query: US-09-041-236-2
Query length: 634
Database: N_Geneseq_36:*
Database sequences: 311585
Database length: 125096042
Search time (sec): 40.470000

score_list:

Sequence	Strd	Orig	ZScore	EScore	Len	Documentation
N_Geneseq_36:X02659	+	3443.00	5419.93	2.7e-294	7475	EP-892047 Seq ID 35. New semap
N_Geneseq_36:X02660	+	3443.00	5418.93	3.1e-294	8192	EP-892047 Seq ID 36. New semap
N_Geneseq_36:X02661	+	3443.00	5420.65	2.5e-294	7000	EP-892047 Seq ID 37. New semap
N_Geneseq_36:X02662	+	3443.00	5420.48	2.5e-294	7108	EP-892047 Seq ID 38. New semap
N_Geneseq_36:X02627	+	3443.00	5431.31	6.3e-295	2636	EP-892047 Seq ID 1. New semaph
N_Geneseq_36:X02658	-	3443.00	5422.60	1.9e-294	5856	EP-892047 Seq ID 34. New semap
N_Geneseq_36:X02666	+	3308.00	5207.12	1.9e-282	6622	EP-892047 Seq ID 42. New semap
N_Geneseq_36:X02665	+	1907.50	2982.40	1.6e-158	8888	EP-892047 Seq ID 41. New semap
N_Geneseq_36:X02628	+	1755.00	2762.41	2.9e-146	1195	EP-892047 Seq ID 2. New semaph
N_Geneseq_36:X02663	+	1069.00	1661.02	6.4e-85	4019	EP-892047 Seq ID 39. New semap
N_Geneseq_36:X02664	+	1049.00	1629.35	3.7e-83	3999	EP-892047 Seq ID 40. New semap
N_Geneseq_36:Q87442	+	649.50	1000.35	4.0e-48	2601	Human semaphorin III cDNA. New
N_Geneseq_36:T60666	+	594.50	906.60	6.7e-43	4391	Mouse CD100 antigen cDNA. Nucl
N_Geneseq_36:V11121	+	591.50	906.89	6.5e-43	2975	Human semaphorin E encoding cD
N_Geneseq_36:X03792	+	591.50	906.88	6.5e-43	2769	Human semaphorin E encoding cDNA
N_Geneseq_36:T60665	+	570.00	869.13	8.2e-41	4157	Human CD100 antigen cDNA. Nucl
N_Geneseq_36:X35367	+	566.00	866.72	1.1e-40	2898	Human semaphorin encoding cDNA
N_Geneseq_36:Q92331	+	547.00	843.91	2.1e-39	1481	Human collapsin cDNA. New anti
N_Geneseq_36:V07279	+	489.50	741.84	1.0e-33	4008	Rat semaphorin W encoding cDNA
N_Geneseq_36:V07280	+	489.50	747.75	4.8e-34	2331	Rat semaphorin W encoding cDNA
N_Geneseq_36:Q87441	+	449.00	681.30	2.4e-30	2854	Grasshopper semaphorin I cDNA.
N_Geneseq_36:V44295	+	432.00	651.95	1.0e-28	3550	Human secreted protein clone C
N_Geneseq_36:V07281	+	420.00	637.59	6.5e-28	2315	Rat semaphorin W encoding cD
N_Geneseq_36:V07282	+	420.00	640.57	4.4e-28	1761	Rat semaphorin W encoding cD
N_Geneseq_36:087446	+	408.00	617.70	8.3e-27	2504	Tribolium semaphorin I cDNA. N
N_Geneseq_36:087445	+	406.00	613.82	1.4e-26	2670	Drosophila semaphorin II cDNA.
N_Geneseq_36:087444	+	399.00	599.58	8.5e-26	3560	Drosophila semaphorin I cDNA.
N_Geneseq_36:T72108	+	382.50	573.52	2.4e-24	3524	Human semaphorin Z gene. Semap
N_Geneseq_36:V28915	+	381.50	572.22	2.8e-24	3432	Human semaphorin Y encoding cD
N_Geneseq_36:V28916	+	381.50	574.48	2.1e-24	2790	Human semaphorin Y encoding cD
N_Geneseq_36:T72107	+	371.50	555.56	2.4e-23	3692	Rat semaphorin Z gene. Semapho
N_Geneseq_36:Q87443	+	364.00	554.79	2.7e-23	1332	Vaccnia virus semaphorin IV c
N_Geneseq_36:V28913	+	357.00	534.14	3.8e-22	3195	Rat semaphorin Y encoding cDNA
N_Geneseq_36:V28914	+	357.00	535.63	3.1e-22	2787	Rat semaphorin Y encoding cDNA
N_Geneseq_36:Q10448	+	124.00	165.77	0.1240	2856	Human basic fibroblast growth
N_Geneseq_36:020914	+	124.00	164.10	0.1536	3328	Human bFGF receptor sequence.
N_Geneseq_36:027658	+	124.00	163.48	0.1665	3525	N-sam cDNA. Gene prod of N-sam
N_Geneseq_36:087447	+	117.50	177.83	0.0284	368	Variola major virus semaphorin
N_Geneseq_36:013311	+	116.00	154.67	0.5722	2469	Basic FGF receptor. Extracellu
N_Geneseq_36:021003	+	116.00	153.85	0.5722	2662	flg receptor protein gene. New
N_Geneseq_36:T31051	+	114.00	150.39	0.8917	2733	Human fibroblast growth factor
N_Geneseq_36:N90388	+	109.50	135.19	6.26	5719	cDNA encoding human platele
N_Geneseq_36:Q60748	+	109.00	167.93	0.0940	265	Human brain Expressed Sequen
N_Geneseq_36:065674	+	109.00	133.89	7.41	5994	Sequence encoding beta subu
N_Geneseq_36:V33392	-	109.00	131.04	10.68	7784	ATP binding cassette transp

seq_name: N_Geneseq_36:X02659

seq_documentation_block:
```
ID   X02659 standard; DNA; 7475 BP.
AC   X02659;
DT   07-MAY-1999 (first entry)
DE   EP-892047 Seq ID 35.
KW   Semaphorin L; human; immunosuppressant; anti-inflammatory; gene therapy;
KW   organ transplantation; inflammation therapy; immunotherapy; agonist;
KW   immunomodulatory; antagonist; ss.
OS   Homo sapiens.
PN   EP-892047-A2.
PD   20-JAN-1999.
PF   06-JUL-1998; 112470.
PR   11-FEB-1998; DE-005371.
PR   09-JUL-1997; DE-029211.
PA   (HMRI) HOECHST MARION ROUSSEL DEUT GMBH.
PI   Ensser A, Fleckenstein B;
DR   WPI; 99-083564/08.
PT   New semaphorin L proteins - used as immunosuppressants and
PT   antiinflammatory agents in organ transplants, inflammation therapy,
PT   immunotherapy and gene therapy
PS   Disclosure; Page 81-85; 135pp; German.
CC   This invention describes a novel human semaphorin L protein. This protein
CC   or its encoding DNA are useful as immunosuppressants and/or
CC   anti-inflammatory agents in organ transplantation, inflammation therapy,
CC   immunotherapy and gene therapy. The DNA can be used to produce knock-out
CC   or knock-in animals for research purposes. The proteins or DNA can be
CC   used to search for the corresponding receptors or to screen for
CC   immunomodulatory agonists or antagonists.
SQ   Sequence  7475 BP;   1667 A;   2073 C;   1974 G;   1761 T;
```

alignment_scores:
```
        Quality: 3443.00              Length:   634
        Ratio:   5.431               Gaps:      0
Percent Similarity: 100.000   Percent Identity: 100.000
```

alignment_block:
US-09-041-236-2 x X02659

Align seg 1/1 to: X02659 from: 1 to: 7475

```
   1 LeuLeuLeuLeuTrpAlaAlaAlaAlaAlaSerAlaGlnGlyHisLeuAr  17
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1064 CTGCTGCTGCTGTGGGCGGCGGCGGCCGCGCGCTCCGCCCAGGGCCACCTAAG 1113

  17 gSerGlyProArgIlePheAlaValTrpGlyHisValGlyGlnAspA  34
     |||||||||||||||||||||||||||||||||||||||||||||||
1114 GAGCGGCGACCCCGCATCTTCGCCGTCTGGAAAGGCCATGTAGGGCAGGACC 1163

  34 rgValAspPheGlyGlnThrGluProHisThrValLeuPheHisGluPro  50
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1164 GGGTGGACTTTGGCCAGACTGAGCCGCCACACCGTGCTTTTCCACGAGCCA 1213

  51 GlySerSerSerValTrpValGlyGlyArgGlyLysValTyrLeuPheAs  67
     |||||||||||||||||||||||||||||||||||||||||||||||||
1214 GGCAGCTCCTCTGTGTGGGTGGGAGGACGTGGCAAGGTCTACCTCTTTGA 1263

  67 pPheProGluGlyLysAsnAlaSerValArgThrValAsnIleGlySerT  84
     |||||||||||||||||||||||||||||||||||||||||||||||||
1264 CTTCCCCGAGGGCAAGAACGCATCTGTGCGCACGGTGAATATCGGCTCCA 1313

  84 hrLysGlySerCysLeuAspLysArgAspCysGluAsnTyrIleThrLeu 100
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1314 CAAAGGGGTCCCTGTCTGGATAAGCGGGACTGCGAGAACTACATCACTCTC 1363

 101 LeuGiuArgArgSerGluGlyLeuLeuAlaCysGlyThrAsnAlaArgH1 117
```

```
1364 |||||||||||||||||||||||||||||||||||||||||||||
     CTGGAGAGGCGGAGTGAGGGGCTGCTGGCCTGTGCCACCAAGCCCGGCA  1413

117  sProSerCysTrpAsnLeuValAlaSerGlyThrValValProLeuGlyGluM  134
1414 CCCCAGCTGCTGGAACCTGGTGAATGGCACTGTGGTGCCACTTGGCGAGA  1463

134  etArgGlyTyrAlaProPheSerProAspGluAsnSerLeuValLeuPhe  150
1464 TGAGAGGCTACGCCCCCTTCAGCCCGGACGAGAACTCCCTGGTTCTGTTT  1513

151  GluGlyAspValTyrSerThrIleArgLysGlnValTyrAsnGlyLy  167
1514 GAAGGGGACGAGGTGTATTCCACCATCCGGAAGCAGGAATACAATGGGAA  1563

167  sIleProArgPheArgGlyArgIleArgGlyGluSerGluGluTyrThrSerA  184
1564 GATCCCTCGGTTCCGCGGCCATCCGGGGCGGGAGAGTGAGCTGTACACCAGTG  1613

184  spThrValMetGlnAsnProGlnPheIleLysAlaThrIleValHisGln  200
1614 ATACTGTCATGCAGAACCCACAGTTCATCAAAGCCACCATCGTGCACCA  1663

201  AspGlnAlaTyrAspLysLysIleTyrTyrPhePheArgGluAspAsnPr  217
1664 GACCAGGCTTACGATGACAAGATCTACTACTTCTTCCGAGAGGACAATCC  1713

217  oAspLysAsnProGluAlaProLeuAsnValSerArgValAlaGlnLeuC  234
1714 TGACAAGAATCCTGAGGCTCCTCTCAAATGTGTCCCGTGTGGCCCAGTTGT  1763

234  ysArgGlyAspGlnGlyGlySerLeuSerSerValSerLysTrpAsn  250
1764 GCAGGGGGACCAGGGTGGGAGAAGTTCACTGTCAGTCTCCAAGTGGAAC  1813

251  ThrPheLeuLysAlaMetLeuValCysSerGlnAlaAlaThrAsnLysAs  267
1814 ACTTTCTGAAAGCCATGCTGGTATGCAGTCAGGCTGCCACCAACAAGAA  1863

267  nPheAsnArgLeuGlnAspValPheLeuLeuProAspProSerGlyGlnT  284
1864 CTTCAACAGGCTGCAAGACGTCTTTCCTGCTCCCTGACCCCAGCGGCCAGT  1913

284  rpArgAspThrArgValTyrGlyValPheSerAsnProTrpAsnTyrSer  300
1914 GGAGGACACCAGGGTCTATGGTGTTTTCTCCAACCCCTGGAACTACTCA  1963

301  AlaValCysTyrSerLeuGlyAspIleAspLysValPheArgPheSe  317
1964 GCCGTCTGTGTGTATTCCCTCGGTGACATTGACAAGGTCTTCCGTACCTC  2013

317  rSerLeuLysTyrHisSerSerLeuProAsnProArgProGlyLysC  334
2014 CTCACTCAAGGGTCTACCACTCAAGCCTTCCCAACCCGCCGGCCTGGCAAGT  2063

334  ysLeuProAspGlnGlnProIleProThrGluProPheGlnValAlaAsp  350
2064 GCCTCCCAGACCAGCAGCCGATACCCACAGAGACCTTCCAGGTGGCTGGAC  2113

351  ArgHisProGluValAlaGlnArgValGluProMetGlyProLeuLysTh  367
2114 CGTCACCCGAGAGGTGGCGCAGAGGGTGGAGCCCATGGGACCCTCGAAGAC  2163

367  rProLeuPheHisSerLysTyrHisTyrGlnLysTyrGlnLeuTyrGlnLy  384
2164 GCCATTGTTCCACTCTAAATACCACTACCAGAAGTGGCCGGTTCACCCGCA  2213

384  etGlnAlaSerHisGlyGluThrPheHisValLeuTyrLeuThrThrAsp  400
2214 TGCAAGCCAGCCACGGGGAGACCTTTCATGTGCTTTACCTAACTACAGAC  2263

401  ArgGlyThrIleHisLysValValGluProGlyGluGlnGluHisSerPh  417
     |||||||||||||||||||||||||||||||||||||||||||||||||
```

```
2264 AGGGCACTATCCACCAAGGTGGTGAACCGGGAGCGGGGAGCAGGAGCACAGCTT  2313

434  eAlaPheAsnIleMetGlyIleGlnProPheArgArgAlaAlaAlaIleG  434
2314 CGCCCTTCAACATCATGGAGATCCAACATTGGAGATCGCCGCGGCTGCCATCC  2363

434  lnThrSerLeuAspPheAlaGluArgArgLysLeuTyrValSerSerGln  450
2364 AGACCATGTCGCTGGATCTGGAGCGGGAGGAAGCGTGTATGTGAGCTCCCAG  2413

451  TrpGluValSerGlnValProLeuAspLeuCysGluValTyrGlyGlyGl  467
2414 TGGGAGGTGTCGAGCCAGGTGCCCCTGGACCTGTGTGAGGTCTATGGCGGGGG  2463

467  yCysHisGlyCysLeuMetSerArgArgAspProTyrCysGlyTrpAspGlnG  484
2464 CTGCCACGGTGCCTCATGTCCCGGAGACCCCTACTGCGGCTGGGACCAGG  2513

484  lyArgCysIleSerIleTyrSerSerGluAspArgSerValLeuGlnSerIle  500
2514 GCCGCTGCATCTCCATCTACAGCTCCGAACGGTCAGTGCTGCAATCCATT  2563

501  AsnProAlaGluProHisLysGluCysProAsnProLysProAspLysSerAl  517
2564 AATCCAGCGGAGCCACACAAGGAGTGTCCCAACCCCAAACCAGACAAGGC  2613

517  aProLeuGlnLysValSerLeuAlaProAsnSerArgArgTyrTyrLeuSerC  534
2614 CCCACTGCAGAAGGTTTCCCTGGCCCCAAACTCTCGCTACTACCTGAGCT  2663

534  ysProMetGluSerArgHisAlaThrTyrSerTrpArgHisLysGluAsn  550
2664 GCCCCATGGAATCCCGCCACGCCACCTACTCATGGCGACACAAGGAGAAC  2713

551  ValGlnInSerCysGluProGlyHisGlnSerProAsnCysIleLeuPh  567
2714 GTGCAGCAGAGCTGCGAACCTGGTCACCAGAGCCCCAACTGCATCCTGTT  2763

567  eIleGluAsnLeuThrAlaGlnInLyrGlyHisTyrPheCysGluAlaG  584
2764 CATCGAGAACCTCACGGCGCAGTACGGCCACTACTTCTGCGAGGCCCC  2813

584  lnGluGlySerTyrPheArgGluAlaGlnInHisTrpGlnLeuProGlu  600
2814 AGGAGGGCTCCTACTTCCGCGAGGCTCAGCACTGGCAGCTGCTGCCCGAG  2863

601  AspGlyIleMetAlaGluHisLeuLeuGlyHisAlaCysAlaLeuAlaAl  617
2864 GACGGCATCATGGCCGAGCACCTGCTGGGTCATGCCTGTGCCCTGGCTGC  2913

617  aSerLeuTrpProLeuGlyValLeuProThrLeuThrLeuLysLeuLeuValH  634
2914 CTCCCTCTGGCCTGGGGGTCGTGCCCACACTCACTCTTGGCTTGGCTGGTCC  2963
```

```
634 is 634
    |
2964·AC 2965
```

PR   09-JUL-1997; DE-029211.
PA   (HMRI) HOECHST MARION ROUSSEL DEUT GMBH.
PI   Ensser A, Fleckenstein B;
DR   WPI: 99-083564/08.
PT   New semaphorin L proteins - used as immunosuppressants and
PT   antiinflammatory agents in organ transplants, inflammation therapy,
PT   immunotherapy and gene therapy
PS   Disclosure; Page 86-91; 135pp; German.
CC   This invention describes a novel human semaphorin L protein. This protein
CC   or its encoding DNA are useful as immunosuppressants and/or
CC   anti-inflammatory agents in organ transplantation, inflammation therapy,
CC   immunotherapy and gene therapy. The DNA can be used to produce knock-out
CC   or knock-in animals for research purposes. The proteins or DNA can be
CC   used to search for the corresponding receptors or to screen for
CC   immunomodulatory agonists or antagonists.
SQ   Sequence 8192 BP; 1839 A; 2313 C; 2177 G; 1863 T;

alignment_scores:
     Quality: 3443.00   Length: 634
     Ratio: 5.431   Gaps: 0
Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
US-09-041-236-2 x X02660

Align seg 1/1 to: X02660 from: 1 to: 8192

    1 LeuLeuLeuLeuTrpAlaAlaAlaAlaAlaSerAlaAlaGlnGlyHisGlyLeuAr   17
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 1064 CTGCTGCTGCTCTGGGCGGCCGCCGCCGCTCCGCCCAGGGCCACCTAAG   1113

   17 gSerGlyProArgIlePheAlaValTrpLysGlyHisValGlyGlnAspA   34
      |||||||||||||||||||||||||||||||||||||||||||||||||||
 1114 GAGCGGACCCCGCCATCTTCGCCGTCTGGAAAGGCCATGTAGGGCAGGACC   1163

   34 rgValAspPheGlyGlnThrGlyHisGluProHisThrValPheHisGluPro   50
      ||||||||||||||||||||||||||||||||||||||||||||||||||||
 1164 GGGTGGACTTTGGCCAGACTGAGCCGCACACGGTGCTTTTCCACGAGCCA   1213

   51 GlySerSerSerValTrpValGlyArgGlyLysValTyrLeuPheAs   67
      ||||||||||||||||||||||||||||||||||||||||||||||||
 1214 GGCAGCTCCTCCGTCTGGGTGGGAGGACGTGGCAAGGTCTACCTCTTTGA   1263

   67 pPheProGluGlyLysAsnAlaSerValArgThrValAsnIleGlySerT   84
      |||||||||||||||||||||||||||||||||||||||||||||||||
 1264 CTTCCCGGAGGGCAAGAACGCATCTGTGCCACGGTGAATATCGGCTCCA   1313

   84 hrLysGlySerCysLeuAspLysSerArgAspCysGluAsnTyrIleThrLeu   100
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1314 CAAAGGGGTCCTGCTTGGATAAGCCGGACTGCGAGAACTACATCACTCTC   1363

  101 LeuGluArgArgSerCysGluGlyLeuLeuAlaCysGlyThrAsnAlaArgH   117
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1364 CTGGAGAGGCGAGCGTGATTCCACCAATCCGGAACGCAACGCCCGGCA   1413

  117 sProSerCysTrpAsnLeuValAsnGlyThrValProLeuGlyGluM   134
      |||||||||||||||||||||||||||||||||||||||||||||||
 1414 CCCCAGCTGCTGGAACCTGGTGAATGGCACTGTGGCCACTTGGCGAGA   1463

  134 etArgGlyTyrAlaProPheSerProGluAsnSerLeuValLeuPhe   150
      |||||||||||||||||||||||||||||||||||||||||||||||
 1464 TGAGAGGCTACGCGCCCCTTCAGCCCGGACGAGAACTCCCTGGTTCTGTTT   1513

  151 GluGlyAspAspGluValTyrSerThrIleArgLysGlnGluTyrAsnGlyLy   167
      |||||||||||||||||||||||||||||||||||||||||||||||||||
 1514 GAAGGGGACGAGGTGTATTCCACCATCCGGAAGCAGGAATACAATGGGAA   1563

  167 sIleProArgPheArgIleGlySerGluGluTyrThrSerA   184
      ||||||||||||||||||||||||||||||||||||||||
 1564 GATCCCTCGGTTCCGCCGCCATCCGGCGAGTGAGCTGTACACCAGTG   1613

  184 spThrValMetGlnAsnProGlnPheIleLysSerAlaThrIleValHisGln   200
      ||||||||||||||||||||||||||||||||||||||||||||||||||||

 1614 ATACTGTCATGCAGAACCCACAGTTCATCAAAGCCACCATCGTGCACCAA   1663
      |||||||||||||||||||||||||||||||||||||||||||||||||||

  201 AspGlnAlaTyrAspAspLysIleTyrPheArgGluAspAsnPr   217
      |||||||||||||||||||||||||||||||||||||||||||||
 1664 GACCAGGCTTACGATGACAAGATCTACTTCCGAGAGGACAATCC   1713

  217 oAspLysAsnProGluAlaProLeuAsnValSerArgValAlaGlnLeuC   234
      |||||||||||||||||||||||||||||||||||||||||||||||||||
 1714 TGACAAGAATCCTGAGGCTCCTCAATGTGTCCCGTGTCCCGTGGCCCAGTTGT 1763

  234 ysArgGlyAspGlnGlyGlyGluSerSerValSerLysTrpAsn   250
      |||||||||||||||||||||||||||||||||||||||||||||
 1764 GCAGGGGGACCAGGGTGGGGAAAGTTCACTGTCAGTCTCCAGTGGAAC   1813

  251 ThrPheLeuLysAlaMetLeuValCysSerAspAlaAlaThrAsnLysAs   267
      |||||||||||||||||||||||||||||||||||||||||||||||||||
 1814 ACTTTTCTGAAAGCCATGCTGGTATGCAGTGATGCTGCCACCAACAAGAA   1863

  267 nPheAsnArgLeuGlnAspValPheLeuLeuPheAspProSerGlyGlnT   284
      |||||||||||||||||||||||||||||||||||||||||||||||||
 1864 CTTCAACAGGCTGCAAGACGTCTTCCTGCTCTTCGATCCCTCCAGCGGCCAGT 1913

  284 rpArgAspThrArgValTyrGlyValPheSerAsnProTrpAsnTyrSer   300
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1914 GGAGGGACACCAGGGTCTATGGTGTTTTCTCCAACCCCTGGACTACTCA   1963

  301 AlaValCysValTyrSerLeuGlyAspIleAspLysValPheArgThrSe   317
      |||||||||||||||||||||||||||||||||||||||||||||||||
 1964 GCCGTCTTGTGTATTCCCTCGGTGACATTGACAAGGTCTTCCGTACCTC   2013

  317 rSerLeuLysGlyTyrHisSerSerLeuProAsnProArgProGlyLysC   334
      |||||||||||||||||||||||||||||||||||||||||||||||||
 2014 CTCACTCAAGGGCTACCACTCAAGCCTTCCCAACCCCGGGCCTGGCAAGT   2063

  334 ysLeuProAspGlnGlnProIleProThrArgGluThrPheGlnValAlaAsp 350
      |||||||||||||||||||||||||||||||||||||||||||||||||||
 2064 GCCTCCCAGACCAGCAGCCGATACCCACCAGAGACCTTCCAGGTGGCTGAC   2113

  351 ArgHisProGluValAlaGlnArgValGluProMetGlyProLeuLysTh   367
      |||||||||||||||||||||||||||||||||||||||||||||||||
 2114 CGTCACCCAGAGGTGGCGCAGAGGGTGGAGCCCATGGGGCCTCTGAAGAC   2163

  367 rProLeuPheHisSerLysTyrHisTyrGlnLysValAlaValHisArgM   384
      |||||||||||||||||||||||||||||||||||||||||||||||||
 2164 GCCATTGTTCCACTCTAAATACCACTACCAGAAAGTGGCCGTTCACCGCA   2213

  384 etGlnAlaSerHisGlyThrPheHisValLeuTyrLeuThrThrAsp   400
      |||||||||||||||||||||||||||||||||||||||||||||||
 2214 TGCAAGCCAGCCACGGGACCTTTCATGTGTTTACCTAACTACACAGAC   2263

  401 ArgGlyThrIleHisLysValValGluProGlyGluGlnHisSerPh   417
      |||||||||||||||||||||||||||||||||||||||||||||||
 2264 AGGGGCACTATCCACAAGGTGGTGGAACCGGGGAGGAGCACCAGCTT   2313

  417 eAlaPheAsnIleMetGluIleGlnProHeArgArgAlaAlaIleG   434
      |||||||||||||||||||||||||||||||||||||||||||||
 2314 CGCCTTCAACATCATGGAGATCCAGCCCTTGCGCGGGCTGCCATCC   2363

  434 lnThrMetSerLeuAspAlaGluArgLysLeuTyrValSerSerGln   450
      |||||||||||||||||||||||||||||||||||||||||||||||
 2364 AGACCATGTCGCTGGATGCTGAGCGGAAGCTGTATGTGAGCTCCCAG   2413

  451 TrpGluValSerGlnValProLeuAspLeuCysGluValTyrGlyGlyl   467
      ||||||||||||||||||||||||||||||||||||||||||||||||
 2414 TGGGAGGTGAGCCAGGTGCCCCTGGACCTGTGTGAGGTCTATGGCGGGGG   2463

  467 yCysHisGlyCysLeuMetSerArgAspProTyrCysGlyTrpAspGlnG   484
      |||||||||||||||||||||||||||||||||||||||||||||||||
 2464 CTGCCACGGTTGCCTCATGTCCCGAGACCCCTACTGCGGCTGGGACCAGG   2513

  484 lyArgCysIleSerIleTyrArgSerGluArgSerValLeuGlnSerIle   500
      ||||||||||||||||||||||||||||||||||||||||||||||||||

```
                Ratio:   5.431                           Gaps:        0

Percent Similarity: 100.000         Percent Identity: 100.000

alignment_block:
US-09-041-236-2 x X02661

Align seg 1/1  to: X02661     from: 1     to: 7000


    1 LeuLeuLeuLeuTrpAlaAlaAlaAlaSerAlaGlnGlyHisLeuAr   17
      |||||||||||||||||||||||||||||||||||||||||||||||
  643 CTGCTCCTGCTCTGGGCGGCCGCCGCCGCCTCCGCCCAGGGCCACCTAAG  692

   17 gSerGlyProArgIlePheAlaValTrpLysGlyHisValGlyGlnAspA   34
      |||||||||||||||||||||||||||||||||||||||||||||||||
  693 GAGCGGAACCCCGACATCTTCGCGTCTGGAAAGGCCATGTAGGGCAGGACC  742

   34 rgValAspPheGlyGlnThrGluProHisThrValPheHisGluPro     50
      |||||||||||||||||||||||||||||||||||||||||||||||
  743 GGGTGGACTTTGGCCAGACTGAGCCGCACAGGGTGCTTTTCCAGAGCCA   792

   51 GlySerSerValTrpValGlyValGlyArgGlyLysValTyrLeuPheas   67
      |||||||||||||||||||||||||||||||||||||||||||||||||
  793 GGCAGCTCCTGTGGTGGGTGGGAGGACGTGGCAAGGTCTACCTCTTTGA   842

   67 pPheProGluGlyLysAsnAlaSerValArgThrValAsnIleGlySerT   84
      |||||||||||||||||||||||||||||||||||||||||||||||||
  843 CTTCCCCGAGGGCAAGAACGCATCTGTGCCACGGTGAATATCGGCTCCA   892

   84 hrLysGlySerCysLeuAspLysArgAspCysGluAsnTyrIleThrLeu  100
      |||||||||||||||||||||||||||||||||||||||||||||||||
  893 CAAAGGGGTCCTGTCTGGATAAGCGGGACTACGAGAACTACATCACTCTC  942

  101 LeuGluArgArgSerGluGlyLeuLeuAlaCysGlyThrAsnAlaArgH1  117
      |||||||||||||||||||||||||||||||||||||||||||||||||
  943 CTGGAGAGGCGGAGTGAGGGGCTGCTGGCCTGTGGCACCAACGCCCGGCA  992

  117 sProSerCysTrpAsnLeuValAsnGlyThrValValProLeuGlyGluM  134
      |||||||||||||||||||||||||||||||||||||||||||||||||
  993 CCCCAGCTGCTGGAACCTGGTGAATGGCACTGTGGTGCCACTTGGCGAGA 1042

  134 etArgGlyTyrAlaProPheSerProAspGluAsnSerLeuValLeuPhe  150
      |||||||||||||||||||||||||||||||||||||||||||||||||
 1043 TGAGAGGTACGGCCCCCCTTCAGCCCGACGAGAACTCCCTGGTTCTGTTT 1092

  151 GluGlyAspGlyuValTyrSerThrIleArgLysGlyGluTyrAsnGlyLy  167
      |||||||||||||||||||||||||||||||||||||||||||||||||
 1093 GAAGGGGACGAGGGTGTATTCCACCATCCGGAAGCAGGAATACAATGGGAA 1142

  167 sIleProArgPheArgIleArgGlySerGluGlyTyrThrSerA        184
      |||||||||||||||||||||||||||||||||||||||||||||||
 1143 GATCCCTCGGTTCCGCCGCATCCGGGGCGAGAGTGAGCTGTACACCAGTG 1192

  184 spThrValMetGlnAsnProGlnPheIleLysAlaThrIleValHisGln  200
      |||||||||||||||||||||||||||||||||||||||||||||||||
 1193 ATACTGTCATGCAGAACCCACAGTTCATCAAAGCCACCATCGTGCACCAA 1242

  201 AspGlnAlaTyrAspAspPheLysIleTyrArgPhePheArgGluAspAsnPr  217
      |||||||||||||||||||||||||||||||||||||||||||||||||
 1243 GACCAGGTTACGATGACAAGATCTACACTTCTTCCGAGAGGACAAATCC 1292

  217 oAspLysAsnProGluAlaProLeuAsnValSerArgValAlaGlnLeuC  234
      |||||||||||||||||||||||||||||||||||||||||||||||||
 1293 TGACAAGAATCCTGAGGCCTCCTCAATCGTCCCGTGTGGCCCCAGTTGT 1342

  234 ysArgGlyAspGlnGlyGlySerLeuSerLeuSerLysTrpAsn        250
      |||||||||||||||||||||||||||||||||||||||||||||||
 1343 GCAGGGGGACCCAGGTGGGAAAGTTCACTGTCAGTCTCCAAGTGGAAC 1392

  251 ThrPheLeuLysAlaMetLeuValCysSerAspAlaAlaThrAsnLysAs  267
      |||||||||||||||||||||||||||||||||||||||||||||||||
 1393 ACTTTTCTGAAAGCCATGCTGGTATGCAGTGAATGCTGCCACCAACAAGAA 1442

  267 nPheAsnArgLeuLeuAspValPheLeuLeuProAspProSerProGlyGlnT  284
```

```
 2514 GCCGCTGCATCCCATTCTACAGTCCGGAACGGTCAGTGCTGCAATCATT 2563
      |||||||||||||||||||||||||||||||||||||||||||||||||
  501 AsnProAlaGluProHisLysGluCysProAsnProLysProAspLysAl  517

 2564 AATCCAGCCGAGCCACCAAGAGAGTGTCCCAACCCAAACCAGACAAGGC 2613
      |||||||||||||||||||||||||||||||||||||||||||||||||
  517 aProLeuGlnLysValSerLeuAlaProAsnSerArgTyrTyrLeuSerC  534

 2614 CCCACTGCAGAAGGTTCCCTGGCCCCAAACTCTCGCTACTACCTGAGCT 2663
      |||||||||||||||||||||||||||||||||||||||||||||||||
  534 ysProMetGluSerProArgHisAlaThrTyrSerThrArgHisLysGluAsn  550

 2664 GCCCCATGGAATCCCGCCACGCCACCTACTCATGGCGCCACAGGAGAAC 2713
      |||||||||||||||||||||||||||||||||||||||||||||||||
  551 ValGluGlnSerCysGluProGlyHisGlnSerProAsnCysIleLeuPh  567

 2714 GTGGAGCAGAGCTGCGAACCTGGTCACCAGAGCCCAACTGCATCCTGTT 2763
      |||||||||||||||||||||||||||||||||||||||||||||||||
  567 eIleGluAsnLeuThrAlaGlnInTyrGlyHisTyrPheCysGluAlaG  584

 2764 CATCGAGAACCTCACGGCGCAGCAGTACGGGCACTACTTCTGCGAGGCCC 2813
      |||||||||||||||||||||||||||||||||||||||||||||||||
  584 lnGluGlySerTyrPheArgGluAlaGlnHisTrpGlnLeuLeuProGlu  600

 2814 AGGAGGGCTCCTACTTCCGCGAGGCTCAGCACTGGCAGCTGCTGCCCGAG 2863
      |||||||||||||||||||||||||||||||||||||||||||||||||
  601 AspGlyIleMetAlaGluHisLeuLeuGlyHisAlaCysAlaLeuAlaAl  617

 2864 GACGGCCATCATGCCGGAGCACCTGGTCATGCCGTCATGCCCTGGCTGC 2913
      |||||||||||||||||||||||||||||||||||||||||||||||||
  617 aSerLeuTrpLeuGlyValLeuProThrLeuThrLeuGlyLeuLeuValH  634

 2914 CTCCCTCTGGCTGGGGTGCTGCCCCACTCACTCTTGGCTTGCTGGTCC 2963
      |||||||||||||||||||||||||||||||||||||||||||||||||
  634 aSerLeuTrpLeuGlyValLeuProThrLeuThrLeuGlyLeuLeuValH  634

  634 1s 634
      ||
 2964 AC 2965
```

seq_name: N_Geneseq_36:X02661

seq_documentation_block:
ID X02661 standard; DNA; 7000 BP.
AC X02661;
DT 07-MAY-1999 (first entry)
DE EP-892047 Seq ID 37.
KW Semaphorin L; human; immunosuppressant; anti-inflammatory; gene therapy;
KW organ transplantation; inflammation therapy; immunotherapy; agonist;
KW immunomodulatory; antagonist; ss.
OS Homo sapiens.
PN EP-892047-A2.
PD 20-JAN-1999.
PF 06-JUL-1998; 112470.
PR 11-FEB-1998; DE-005371.
PR 09-JUL-1997; DE-029211.
PA (HMRI) HOECHST MARION ROUSSEL DEUT GMBH.
PI Ensser A, Fleckenstein B;
DR WPI; 99-083564/08.
PT New semaphorin L proteins - used as immunosuppressants and
PT antiinflammatory agents in organ transplants, inflammation therapy,
PT immunotherapy and gene therapy
PS Disclosure; Page 91-95; 135pp; German.
CC This invention describes a novel human semaphorin L protein. This protein
CC or its encoding DNA are useful as immunosuppressants and/or
CC anti-inflammatory agents in organ transplantation, inflammation therapy,
CC immunotherapy and gene therapy. The DNA can be used to produce knock-out
CC or knock-in animals for research purposes. The proteins or DNA can be
CC used to search for the corresponding receptors or to screen for
CC immunomodulatory agonists or antagonists.
SQ Sequence 7000 BP; 1586 A; 1940 C; 1850 G; 1624 T;

alignment_scores:
      Quality: 3443.00   Length: 634

```
2343 CATCGAGAACCTCACGGCGCAGCAGTACGGCCACTACTTCTGCGAGGCCC 2392
                                                       |||||||||||||||||||||||||||||||||||||||||||||||||||
 584 lnGluGlySerTyrPheArgGluAlaGlnHisTrpGlnLeuLeuProGlu 600
2393 AGGAGGGCTCCTACTTCCGCGAGGCTCAGCACTGGCAGCTGCTGCCCGAG 2442

 601 AspGlyIleMetAlaGluHisLeuLeuGlyHisAlaCysAlaLeuAlaAl 617
     ||||||||||||||||||||||||||||||||||||||||||||||||||
2443 GACGGCATCATGGCCGAGCACCTGCTGGGTCATGCCTGCGCCCTGGCTGC 2492

 617 aSerLeuTrpLeuGlyValLeuProThrLeuThrLeuLeuLeuValH 634
     |||||||||||||||||||||||||||||||||||||||||||||||
2493 CTCCCTGTGGCTGGGGGTGCTGCCCACACTCACTCTTGGCTTGCTGGTCC 2542

634 is 634
      ||
2543 AC 2544
```

seq_name: N_Geneseq_36:X02662

```
seq_documentation_block:
ID   X02662 standard; DNA; 7108 BP.
AC   X02662;
DT   07-MAY-1999 (first entry)
DE   EP-892047 Seq ID 38.
KW   Semaphorin L; human; immunosuppressant; anti-inflammatory; gene therapy;
KW   organ transplantation; inflammation therapy; immunotherapy; agonist;
KW   immunomodulatory; antagonist; ss.
OS   Homo sapiens.
PN   EP-892047-A2.
PD   20-JAN-1999.
PF   06-JUL-1998; 112470.
PR   11-FEB-1998; DE-005371.
PR   09-JUL-1997; DE-029211.
PA   (HMRI ) HOECHST MARION ROUSSEL DEUT GMBH.
PI   Enser A, Fleckenstein B;
DR   WPI; 99-083564/08.
PT   New semaphorin L proteins - used as immunosuppressants and
PT   antiinflammatory agents in organ transplants, inflammation therapy,
PT   immunotherapy and gene therapy
PS   Disclosure. Page 96-100; 135pp; German.
CC   This invention describes a novel human semaphorin L protein. This protein
CC   or its encoding DNA are useful as immunosuppressants and/or
CC   anti-inflammatory agents in organ transplantation, inflammation therapy,
CC   immunotherapy and gene therapy. The DNA can be used to produce knock-out
CC   or knock-in animals for research purposes. The proteins or DNA can be
CC   used to search for the corresponding receptors or to screen for
CC   immunomodulatory agonists or antagonists.
SQ   Sequence 7108 BP;   1617 A;   1970 C;   1875 G;   1646 T;
```

```
alignment_scores:
     Quality: 3443.00              Length:     634
       Ratio:  5.431                 Gaps:       0
Percent Similarity: 100.000    Percent Identity: 100.000
```

```
alignment_block:
US-09-041-236-2 x X02662

Align seq 1/1 to: X02662   from: 1   to: 7108

   1 LeuLeuLeuLeuTrpAlaAlaAlaAlaSerAlaAlaGlyHisLeuAr 17
     |||||||||||||||||||||||||||||||||||||||||||||||
 643 CTGCTGCTGCTGCTGGGCGCGGCGGCCGCCTCCGCGCCAGGGCCACCTAAG 692

  17 gSerGlyProArgIlePheAlaValTrpLysValGlyGlnAspA 34
     ||||||||||||||||||||||||||||||||||||||||||||
 693 GAGCGGACCCCGCCATCTTCGCCGTCTGGAAAGGCCATGTAGGCAGGACC 742

  34 rgValAspPheGlyGlnThrGluProHisThrValLeuPheHisGluPro 50
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 743 GTGGAGGAGCGTCGCAACCTGGTCACCAGAGCCCCACCGGTGCTTTTCCAGGAGCCA 792
```

```
1443 CTTCAACAGCCTGCAAGACGTCTTCCTGCTCCCTGCCCAGCCGCCAGT 1492
     ||||||||||||||||||||||||||||||||||||||||||||||||

 284 rpArgAspThrArgValTyrGlyValPheSerAsnProTrpAsnTyrSer 300
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1493 GGAGGGACACCAGGGTCATGGTGTTTCTCCAACCCCTGGAACTACTCA 1542

 301 AlaValCysValTyrSerLeuGlyAspIleAspLysSerValPheArgThrSe 317
     |||||||||||||||||||||||||||||||||||||||||||||||||
1543 GCCGTCGTGTGTATTCCCTCGGTGACAATGACAAGGTCTTCCGTACCTC 1592

 317 rSerLeuLysGlyTyrHisSerSerLeuProAsnProArgProGlyLysC 334
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1593 CTCACTCAAGGGCTACCACTCAAGCCTTCCAACCCGCGCCTGCCAAGT 1642

 334 ysLeuProAspGlnProThrGluThrPheGlnValAlaAlaAsp 350
     |||||||||||||||||||||||||||||||||||||||||||||
1643 GCCTCCCAGACCCAGCAGCCGATACCCACAGAGACCTTCCAGGTGGCTGAC 1692

 351 ArgHisProGlyValAlaGlnAsnArgValGluProMetGlyProLeuLysTh 367
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1693 CGTCACCCAGAGGTGGCGCAGAGGGTGGAGCCCATGGGGCCTCTGAAGAC 1742

 367 rProLeuPheHisSerLysTyrHisTyrGlnLnLysValAlaValHisArgM 384
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1743 GCCATTGTTCCACTCTAAATACCACTACCAGAAAGTGGCCGTTCACCGCA 1792

 384 etGlnAlaSerHisGlyGluThrPheHisValLeuTyrLeuThrThrAsp 400
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1793 TGCAAGCCAGCCACGGGAGAGACCTTTCATGTGCTTTACCTAACTACAGAC 1842

 401 ArgGlyThrIleHisIleLysSerValGluProGlyGluGlnHisSerPh 417
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1843 AGGGGCACTATCCACAAGTGGTGAACGGGGGAGCAGGAGCACAGCTT 1892

 417 eAlaPheAsnIleMetGluIleGlnProPheArgArgAlaAlaAlaIleG 434
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1893 CGCCTTCAACATCATGGAGATCCAGCCCTTCCGGCGCGCGGCTGCCATCC 1942

 434 lnThrMetSerLeuAspAlaGluArgArgLysLeuTyrValSerSerGln 450
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1943 AGACCATGTCGCTGGATGCTGAGCGGAGAAGCTGTATGTGAGCTCCCAG 1992

 451 TrpGluValSerGlnValProLeuAspLeuCysLeuValTyrGlyGlyGl 467
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1993 TGGGAGGTGAGCCAGGTGCCCCTGGACCTGTGTCTGGTGTATGGCGGGGG 2042

 467 yCysHisGlyCysLeuMetSerArgAspProTyrCysGlyTrpAspGlnG 484
     ||||||||||||||||||||||||||||||||||||||||||||||||||
2043 CTGCCACGGGTGCCTCATGTCCCGAGACCCCTACTGCGGCTGGGACCAGG 2092

 484 lyArgCysIleSerIleTyrSerSerGluArgSerValLeuGlnSerIle 500
     ||||||||||||||||||||||||||||||||||||||||||||||||||
2093 GCCGGCTGCATCTCCATCTACAGCTCCGAACGGGTCAGTGCTGCAATCCATT 2142

 501 AsnProAlaGluProHisLysGluCysProAsnProLysProAspLysAl 517
     ||||||||||||||||||||||||||||||||||||||||||||||||||
2143 AATCCAGCCGAGCCACACAAGGAGTGTCCCAACCCCAAACCAGACAAGGC 2192

 517 aProLeuGlnLysValSerLeuAlaProAsnSerArgTyrTyrLeuSerC 534
     ||||||||||||||||||||||||||||||||||||||||||||||||||
2193 CCCACTGCAGAAGGTTTCCCTGGCCCCAAACTCTCGCTACTACCTGAGCT 2242

 534 ysProMetGluSerArgHisAlaThrTyrSerTrpArgHisLysGluAsn 550
     ||||||||||||||||||||||||||||||||||||||||||||||||||
2243 GCCCCATGGAATCCCGCCACGCCACCTACTCATGGCGCCACAAGGAGAAC 2292

 551 ValGluGlnSerCysGluProGlyHisGlnSerProAsnCysIleLeuPh 567
     ||||||||||||||||||||||||||||||||||||||||||||||||||
2293 GTGGAGCAGAGCTGCGAACCTGGTCACCAGAGCCCCAACTGCATCCTGTT 2342

 567 eIleGluAsnLeuThrArgAlaGlnInTyrGlyHisTyrPheCysGluAlaG 584
     ||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
  51  GlySerSerValTrpValGlyGlyArgGlyLysValTyrLeuPheAs   67
      ||||||||||||||||||||||||||||||||||||||||||||||||
 793  GGCAGCTCCTGTGTGGGTGGGAGGACGTGGCCAAGGTCTACCTCTTTGA  842

  67  pPheProGluGlyLysAsnAlaSerValArgThrValAsnIleGlySerT   84
      |||||||||||||||||||||||||||||||||||||||||||||||||
 843  CTTCCCCAGGGGCAAGAACGCATCTGTGCGCACGGTGAATATCGGCTCCA  892

  84  hrLysGlySerCysLeuAspLysArgAspCysGlyAsnTyrIleThrLeu  100
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 893  CAAAGGGGTCCTGTCTGATAAGCGGACTGCGAGAACTACATCACTCTC    942

 101  LeuGluArgGluSerGluGlyLeuLeuAlaCysGlyTyrThrAsnAlaArgHi  117
      |||||||||||||||||||||||||||||||||||||||||||||||||||
 943  CTGGAGAGGCGGAGTGAGGGGCTGCTGGCCTGTGGCCACCAACCGCCCGGCA  992

 117  sProSerCysTrpAsnLeuVaLAsnGlyThrValValProLeuGlyLeuM  134
      |||||||||||||||||||||||||||||||||||||||||||||||||
 993  CCCCAGCTGCTGGAACCTGGTGAATGGCACTGTGGTGCCACTTGGCCAGA  1042

 134  etArgGlyTyrAlaProPheSerProAspGluAsnSerLeuValLeuPhe  150
      ||||||||||||||||||||||||||||||||||||||||||||||||||
1043  TGAGAGGCTACGCCCCCTTCAGCCCGGACGAGAACTCCCTGGTTCTGTTT  1092

 151  GluGlyAspGluValTyrSerThrIleArgLysGlnLeuTyrAsnGlyLy  167
      ||||||||||||||||||||||||||||||||||||||||||||||||||
1093  GAAGGGGACGAGGTGTATTCCACCATCCGAAGCAGGAATACAATGGGAA  1142

 167  sIleProArgPheArgArgIleArgGlyGluSerGluLeuTyrThrSerA  184
      ||||||||||||||||||||||||||||||||||||||||||||||||||
1143  GATCCCCCGGTTCCGCCGCATCCGGGGCGAGAGTGAGCTGTACACCAGTG  1192

 184  spThrValMetGlnAsnProGlnPheIleLysAlaThrIleValHisGln  200
      ||||||||||||||||||||||||||||||||||||||||||||||||||
1193  ATACTGTCATGCAGAACCCACAGTTCATCAAAGCCACCATCGTGCACCAA  1242

 201  AspGlnAlaTyrAspAspLysIleTyrTyrPhePheArgGluAspAsnPr  217
      ||||||||||||||||||||||||||||||||||||||||||||||||||
1243  GACCAGGCTTACGATGACAAGATCTACTACTTCTTCCGAGAGGACAATCC  1292

 217  oAspLysAsnProGluAlaProLeuAsnValSerArgValAlaGlnLeuC  234
      ||||||||||||||||||||||||||||||||||||||||||||||||||
1293  TGACAAGAATCCTGAGGCTCCTCTCAATGTCTCCCGTGTGGCCCAGTTGT  1342

 234  ysArgGlyAspGlyAspGlyGlyGluSerSerValSerLysTrpTrpAsn  250
      ||||||||||||||||||||||||||||||||||||||||||||||||||
1343  GCAGGGGGGACGGGGATGGGGGAAAGTTCACTGTCAGTCTCCAGTGGAAC  1392

 251  ThrPheLeuLysAlaMetLeuValCysSerAspSerAlaAlaThrAsnLysAs  267
      |||||||||||||||||||||||||||||||||||||||||||||||||||
1393  ACTTTTCTGAAAGCCATGCTGGTATGCAGTGATAGCGCTGCCACCAACAAGAA  1442

 267  nPheAsnArgLeuGlnAspValPheLeuLeuProAspSerProSerGlyGlnT  284
      |||||||||||||||||||||||||||||||||||||||||||||||||||
1443  CTTCAACAGGCTGCAAGACGTCTTCCTGCTCCCTGACTCCCCAGCGGCCAGT  1492

 284  rpArgAspThrArgValTyrGlyValPheSerAsnProTrpAsnTyrSer  300
      ||||||||||||||||||||||||||||||||||||||||||||||||||
1493  GGAGGGGACACCAGGGTCTATGGTGTTTTCTCCAACCCCTGGAACTACTCA  1542

 301  AlaValCysValTyrSerLeuGlyAspIleAspLysLysValPheArgThrSe  317
      |||||||||||||||||||||||||||||||||||||||||||||||||
1543  GCCGTCTGTGTGTATTCCCTGGGTGACATTGACAAGGTCTTCCGTACCTC  1592

 317  rSerLeuGlyTyrHisSerSerLeuProAsnProArgProGlyLysC  334
      |||||||||||||||||||||||||||||||||||||||||||||||
1593  CTCACTCAAGGGCTACCACTCAAGCCTTCCCAACCCCCGGCCGGGCAAGT  1642

 334  ysLeuProAspGlnProIleProThrGlyPheGlnValAlaAlaAsp  350
      |||||||||||||||||||||||||||||||||||||||||||||||
1643  GCCTCCCAGACCAGCCGATACCCCACAGGACCTTCCAGGTGGCTGCTGAC  1692

 351  ArgHisProGluValAlaAlaGlnArgValGlyProMetGlyProLeuLysTh  367
```

```
                                                    |||||||||||||||||||||||||||||||||||||||||||||||||  1742
1693  CGTCACCCAGAGGTGGCGCCAGAGGTGGACGGGCCTCTGAAGAC

 367  rProLeuPheHisSerLysTyrHisTyrGlnLysValAlaValAlaValHisArgM  384
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||
1743  GCCATTGTTCCACTCTAAATACCACCTACCAGAAAGTGGCCGTTCACCGCA  1792

 384  etGlnAlaSerHisGlyLeuThrPheHisValLeuTyrLeuThrThrAsp  400
      ||||||||||||||||||||||||||||||||||||||||||||||||||
1793  TGCAAGCCAGCCACGGGAGACCTTCATGTGCTTTACCTAACTACAGAC  1842

 401  ArgGlyThrIleHisLysLysValGluProGlyGlnGlnHisSerProHis  417
      ||||||||||||||||||||||||||||||||||||||||||||||||||
1843  AGGGGCACTATCCACAAGGTGGTGAACCGGGGGACAGCAGGACCACAGCTT  1892

 417  eAlaPheAsnIleMetGluIleGlnProPheArgArgAlaAlaAlaIleG  434
      ||||||||||||||||||||||||||||||||||||||||||||||||||
1893  CGCCTTCAACATCATGGAGATCCAGCCCTTCCGCCGCGCGGCTGCCATCC  1942

 434  lnThrMetSerLeuAspAlaGluArgArgLysLeuTyrValSerSerGln  450
      ||||||||||||||||||||||||||||||||||||||||||||||||||
1943  AGACCATGTCGCTGGATGCTGAGCGGAGGAAGCTGTATGTGAGCTCCCAG  1992

 451  TrpGluValSerGlnValProLeuAspLeuCysGluValTyrGlyValGl  467
      ||||||||||||||||||||||||||||||||||||||||||||||||||
1993  TGGGAGGTGAGCCAGGTGCCCCTGGACCTGTGTGAGGTCTATGGCGGGGG  2042

 467  yCysHisGlyCysLeuMetSerArgArgAspProTyrCysGlyTrpProAspGlnG  484
      |||||||||||||||||||||||||||||||||||||||||||||||||||||
2043  CTGCCACGGTTGCCTCATGTCCCGAAGCCGAGACCCCTACTGCGGCTGGGACCAGG  2092

 484  lyArgCysIleSerIleTyrSerSerGluArgSerValLeuGlnSerIle  500
      ||||||||||||||||||||||||||||||||||||||||||||||||||
2093  GCCGCTGCATCTCCATCTACAGCTCCGACAGGTCCAGTGCTGCTGCAATCCATT  2142

 501  AsnProAlaGluProHisLysGluCysProAsnProLysProAspProLysSerA  517
      |||||||||||||||||||||||||||||||||||||||||||||||||||||
2143  AATCCAGCGGAGCCGCACAAGGAGTGTCCAACCCCAAACCAGACAAGGC  2192

 517  aProLeuGlnLysValSerArgLeuAlaProAsnSerArgTyrTyrLeuSerC  534
      |||||||||||||||||||||||||||||||||||||||||||||||||||
2193  CCCACTGCAGAAGGTTTCCCGTCTGGCCCCAAACTCTCGGCTACTACCTGAGCT  2242

 534  ysProMetGluSerArgHisAlaThrTyrSerTrpArgHisTyrLysGluAsn  550
      |||||||||||||||||||||||||||||||||||||||||||||||||||
2243  GCCCCATGGAGAATCCCGGCCACACGCCCACCTACTCATGGCGCCACAAGGAGAAC  2292

 551  ValGluGlnSerCysGluProGlyHisGlnSerProAsnCysIleLeuPh  567
      ||||||||||||||||||||||||||||||||||||||||||||||||||
2293  GTGGAGCAGAGCTGCGAACCTGGTCACCAGAGCCCCAACTGCATCCTGTT  2342

 567  eIleGluAsnLeuThrAlaGlnIleAsnTyrGlyHisTyrPheCysGluAlaG  584
      |||||||||||||||||||||||||||||||||||||||||||||||||||
2343  CATCGAGAACCTCACGGCGCAGATACGGCCCACTACTTCTGCGAGGCCC  2392

 584  lnGluGlySerTyrPheArgGluArgGlnHisHisTrpLeuArgProGlu  600
      ||||||||||||||||||||||||||||||||||||||||||||||||||
2393  AGGAGGGCTCCTACTTCCGGGAGCGGCAGCACCACTGGCTGCGCCCCGAG  2442

 601  AspGlyIleMetAlaGluHisHisLeuLeuGlyHisAlaCysAlaLeuAlaAl  617
      |||||||||||||||||||||||||||||||||||||||||||||||||
2443  GACGGCATCATGGCCGAGCACCTGCTGGGTCATGCCTGTGCCCTGGCTGC  2492

 617  aSerLeuTrpLeuValLeuProThrLeuGlyLeuLeuLeuLeuValH  634
      ||||||||||||||||||||||||||||||||||||||||||||||
2493  CTCCCTCTGGCTGGTGGGGCTGCCCACCCTACTCTTGGCTTGCTGGTCC  2542

 634  is 634
      ||
2543  AC 2544

seq_name: N_Geneseq_36:X02627
```

```
seq_documentation_block:
ID  X02627 standard; DNA; 2636 BP.
AC  X02627;
DT  07-MAY-1999 (first entry)
DE  EP-892047 Seq ID 1.
KW  Semaphorin L; human; immunosuppressant; anti-inflammatory; gene therapy;
KW  organ transplantation; inflammation therapy; immunotherapy; agonist;
KW  immunomodulatory; antagonist; ss.
OS  Homo sapiens.
PN  EP-892047-A2.
PD  20-JAN-1999.
PF  06-JUL-1998; DE-005371.
PR  11-FEB-1998; DE-005371.
PR  09-JUL-1997; DE-029211.
PA  (HMRI ) HOECHST MARION ROUSSEL DEUT GMBH.
PI  Ensser A, Fleckenstein B;
DR  WPI; 99-083564/08.
PT  New semaphorin L proteins - used as immunosuppressants and
PT  antiinflammatory agents in organ transplants, inflammation therapy,
PT  immunotherapy and gene therapy
PS  Disclosure; Page 58-60; 135pp; German.
CC  This invention describes a novel human semaphorin L protein. This protein
CC  or its encoding DNA are useful as immunosuppressants and/or
CC  anti-inflammatory agents in organ transplantation, inflammation therapy,
CC  immunotherapy and gene therapy. The DNA can be used to produce knock-out
CC  or knock-in animals for research purposes. The proteins or DNA can be
CC  used to search for the corresponding receptors or to screen for
CC  immunomodulatory agonists or antagonists.
SQ  Sequence  2636 BP;  536 A;  835 C;  764 G;  501 T;

alignment_scores:
          Quality: 3443.00          Length:      634
            Ratio:    5.431           Gaps:        0
Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-041-236-2 x X02627

Align seq 1/1  to: X02627     from: 1    to: 2636
```

```
217  oAspLysAsnProGluAlaProLeuAsnValSerArgValAlaGlnLeuC  234
     |||||||||||||||||||||||||||||||||||||||||||||||||
1494 TGACAAGAAACCTGAGCCTCTCCAATGTGTCCCGTGTGGCCCAGTTGT   1445

234  ysArgGlyAspGlyGlySerLeuSerValSerLeuTrpAsn          250
     |||||||||||||||||||||||||||||||||||||||||
1444 GCAGGGGGACCACCAGGTGGGGAAAGTTCACTGTCTCCAAGTGGAAC    1395

251  ThrPheLeuLysAlaMetLeuVallCysSerAspAlaAlaThrAsnLysAs 267
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1394 ACTTTTCTGAAAGCCATGCTGGTATGCAGTGATGCTGCCACCAACAAGAA 1345

267  nPheAsnArgLeuGlnAspValPheLeuLeuProAspProSerGlyGlnT 284
     |||||||||||||||||||||||||||||||||||||||||||||||||
1344 CTTCAACAGGCTGCAAGACGTCTTCCTGCTTCCTGACCCCAGCGGCCAGT 1295

284  rPArgAspThrArgValTyrGlyValPheSerAsnProTrpAsnTyrSer 300
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1294 GGAGGGACACCAGGGTCTATGGTGTTTTCTCCAACCCCTGGAACTACTCA 1245

301  AlaValCysValTyrSerLeuGlyAspIleAspIlePheArgArgThrSe 317
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1244 GCCGTCTGTGTGTATTCCCTCGGTGACATTGACAGGTCTTCCGTACCCTC 1195

317  rSerLeuLysGlyTyrHisSerSerLeuProAsnProArgProGlyLysC 334
     |||||||||||||||||||||||||||||||||||||||||||||||||
1194 CTCACTCAAGGGCTACCACTCAAGCCTTCCAACCCGGCCTGGCAAGT    1145

334  ysLeuProAspGlnGlnProIleProThrGluThrPheGlnValAlaAsp 350
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1144 GCCTCCCAACCCAGCAGCCGATACCCACAGAGACCTTCCAGGTGGCTGAC 1095

351  ArgHisProGlyValAlaGlnValAlaGluProMetGlyProLeuLysTh 367
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1094 CGTCACCCAGGGTGGCGCAGAGGTGGCAGCCCATGGGAGCCCTCTGAAGAC 1045

367  rProLeuPheHisSerLysTyrHisTyrGlnLysTyrGlnValAlaValHisArgM 384
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1044 GCCATTGTTCCACTCTAAATACCACTACCAGAAAGTGGCCGTTCACCGCA 995

384  etGlnAlaSerHisGlyGluThrPheHisValLeuTyrLeuThrThrAsp 400
     ||||||||||||||||||||||||||||||||||||||||||||||||||
994  TGCAAGCCAGCCACGGGGAGAACCTTCATGTGCTTTACCTAACTACAGAC 945

401  ArgGlyThrIleHisLysHisValValGluProGlyGluGlnHisSerPh 417
     ||||||||||||||||||||||||||||||||||||||||||||||||||
944  AGGGGCACTATCCACAAGGTGGTGAACCCGGGGGAGCAGGAGCACAGCTT 895

417  eAlaPheAsnIleMetGluIleGlnProPheArgArgAlaAlaAlaIleG 434
     ||||||||||||||||||||||||||||||||||||||||||||||||||
894  CGCCTTCAACATCATGGAGATCCAGCCCTTCCGCGCGCGGCTGCCATCC  845

434  lnThrMetSerLeuAspAlaGluArgArgLysLeuTyrValSerSerGln 450
     ||||||||||||||||||||||||||||||||||||||||||||||||||
844  AGACCATGTCGCTGGATGCCGAGCGGAGGAAGCTGTATGTGAGCTCCCAG 795

451  TrpGluValSerGlnValProLeuAspLeuCysGlyValTyrGlyGlyGl 467
     ||||||||||||||||||||||||||||||||||||||||||||||||||
794  TGGGAGGTGAGCCAGGTGCCCCTGGACCTGTGTGGAGGTCATGGGGGGGG 745

467  yCysHisGlyCysLeuMetSerArgArgAspProTyrCysGlyTrpAspGlnG 484
     ||||||||||||||||||||||||||||||||||||||||||||||||||
744  CTGCCACGGTTGCCTCATGTCCCGGAGACCCCTACTGCGGCTGGGACCAGG 695

484  lyArgCysIleSerIleTyrSerSerGluArgSerValLeuGlnSerIle 500
     ||||||||||||||||||||||||||||||||||||||||||||||||||
694  GCCGGCTGCATCTCCATCTACAGCTCCGAACGGTCAGTGCTGCAATCCATT 645

501  AsnProAlaGluProHisLysCysProAsnProLysProAspLysLysAl 517
     ||||||||||||||||||||||||||||||||||||||||||||||||||
644  AATCCAGCCGAGCCACACAAGAGTGTCCCCAACCCAAACCAGACAAGGC  595

517  aProLeuGlnLysValSerLeuAlaProAsnSerArgTyrTyrLeuSerC 534
```

```
                                                       545
     |||||||||||||||||||||||||||||||||||||||||||||||
594  CCCACTCAGAAGGTTCCCTGGCCCCAAACTCGCTACTACCTGAGCT    545

534  ysProMetGluSerArgHisAlaThrTyrSerTrpArgHisLysGluAsn 550
     ||||||||||||||||||||||||||||||||||||||||||||||||||
544  GCCCCATGGAATCCCCACCCACGCCACCTACTCATGGGCCACAAGGAGAAC 495

551  ValGluGlnSerCysGluProGlyHisGlnSerProAsnCysIleLeuPh 567
     ||||||||||||||||||||||||||||||||||||||||||||||||||
494  GTGGAGCAGAGCTGCGAACCTGGTCACCAGAGCCCAACTGCATCCTGTT  445

567  eIleGluAsnLeuThrAlaGlnInTyrGlyHisTyrPheCysGluAlaG 584
     |||||||||||||||||||||||||||||||||||||||||||||||||
444  CATCGAGAACCTCACGGCGCAGCAGTACGGCCACTACTTCTGCGAGGCCC 395

584  lnGluGlySerTyrPheArgGluAlaGlnHisTrpGlnLeuLeuProGlu 600
     ||||||||||||||||||||||||||||||||||||||||||||||||||
394  AGGAGGGCTCCTACTTCCGCGAGGCTCAGCACTGGCAGCTGCTGCCCGAG 345

601  AspGlyIleMetAlaGluHisLeuLeuGlyHisAlaCysAlaLeuAlaAl 617
     ||||||||||||||||||||||||||||||||||||||||||||||||||
344  GACGGGCATCATGGCCGAGCACCTGCTGGGTCATGCCTGTGCCCTGGCTGC 295

617  aSerLeuTrpLeuGlyValLeuProThrLeuThrLeuLeuGlyLeuLeuValH 634
     ||||||||||||||||||||||||||||||||||||||||||||||||||
294  CTCCCTCTGGCTGGGCGTGCTGCCCACTCTCACTCTTGGCTTGCTGGTCC 245

634  is 634
     ||
244  AC 243
```

seq_name: N_Geneseq_36:X02666

alignment_scores:
```
    Quality:   3308.00                    Length:    607
    Ratio:     5.450                       Gaps:      0
Percent Similarity: 100.000    Percent Identity: 100.000
```

alignment_block:
US-09-041-235-2 x. X02666

align seg 1/1 to: X02666   from: 1   to: 6622

```
 311 AspLysValPheArgThrSerSerLeuLysGlyTyrHisSerSerLeuPr  327
     |||||||||||||||||||||||||||||||||||||||||||||||||
1071 GACAAGGTCTTCCGTACCTCCTCACTCAAGGGCTACCACTCAAGCCTTCC 1120

 327 OAsnProArgProGlyLysCysLeuProAspGlnProIleProHisThrG  344
     |||||||||||||||||||||||||||||||||||||||||||||||||
1121 CAACCCGAGGCCGGCCTGGCAAGTGCCTCCCAGACCCAGCCGATACCCACAG 1170

 344 luThrPheGlnValAlaAspArgHisProGluValAlaGlnArgValGlu  360
     |||||||||||||||||||||||||||||||||||||||||||||||||
1171 AGACCTTCCAGGTGGCTGACCGTCACCCAGAGGTGGCCAGAGGGTGGAG 1220

 361 ProMetGlyProLeuLysThrProLeuPheHisSerSerLysTyrHisTyrGl  377
     |||||||||||||||||||||||||||||||||||||||||||||||||
1221 CCCATGGGGCCCTCTGAAGACGGCCATTGTTCCACTCTAAATACCACTACCA 1270

 377 nLysValAlaValHisArgMetGlnAlaSerHisGlyGlnThrPheHisSv  394
     |||||||||||||||||||||||||||||||||||||||||||||||||
1271 GAAAGTGGCCGTTCACCCGATGCAAGCCAGCCACGGGGAGACCTTTCATG 1320

 394 aLeuTyrLeuThrThrAspArgGlyThrIleHisIleLysLysValGluPro  410
     |||||||||||||||||||||||||||||||||||||||||||||||||
1321 TGCTTTACCTAACTACAGACAGGGCACTATCCACAAGGTGGTGGAACCG 1370

 411 GlyGluGlnHisSerPheAlaPheAsnIleMetGluIleGlnProPh  427
     |||||||||||||||||||||||||||||||||||||||||||||||||
1371 GGGGAGCAGGAGCACCAGCTTCGCCTTCAACATCATGGAGATCCAGCCCTT 1420

 427 eArgArgAlaAlaAlaIleGlnThrMetSerLeuAspAlaGluArgArgL  444
     |||||||||||||||||||||||||||||||||||||||||||||||||
1421 CCGGCCGGCGGCTGCCATCCAGACCATGTCGCTGGATGCTGAGCGGGAGA 1470

 444 ysLeuTyrLeuThrValSerSerGlnTrpGluValSerGlnValProLeuAspLeu  460
     |||||||||||||||||||||||||||||||||||||||||||||||||
1471 AGCTGTATGTGAGCGTCCCAGTGGGAGGTGAGCCAGGTGCCCCTGGACCTG 1520

 461 CysGluValTyrGlyGlyTyrGlyHisGlyTyrCysLeuMetSerArgAspPr  477
     |||||||||||||||||||||||||||||||||||||||||||||||||
1521 TGTGAGGTCTATGGCGGGGGGTGCCACGGTTGCCTCATGTCCCGAGACCC 1570

 477 oTyrCysGlyTyrTrpAspGlnIleArgCysIleSerIleTeTyrSerSerGluA  494
     |||||||||||||||||||||||||||||||||||||||||||||||||
1571 CTACTGCGGCTGGGACCAGATGGCCCGCTGCATCTCCATCTACAGCTCCGAAC 1620

 494 rgSerValLeuGlnSerIleAsnProAlaGluProHisLysGluCysPro  510
     |||||||||||||||||||||||||||||||||||||||||||||||||
1621 GGTCAGTGCTGCAATCCATTAATCCAGCCGAGCCCACACAAGGAGTGTCCC 1670

 511 AsnProLysProAspLysProLeuGlnLysValSerLeuAlaProAs  527
     |||||||||||||||||||||||||||||||||||||||||||||||||
1671 AACCCCAAACCAGACAAGCCCCACTGCAGAAGGTTTCCCTGGCCCCAAA 1720

 527 nSerArgTyrTyrLeuSerCysProMetGluSerArgHisAlaThrTyrS  544
     |||||||||||||||||||||||||||||||||||||||||||||||||
1721 CTCTCGCTACTACCTGAGCTGCCCCATGGAGAATCCCGCCACGCCACCTACT 1770

 544 erTrpArgHisLysGluAsnValGluGlnSerCysProGlyHisGln  560
     |||||||||||||||||||||||||||||||||||||||||||||||||
1771 CATGGCGCCACAGGAGACGTGGAGCAGAGCTGCGAACCTGGTCACCAG 1820

 561 SerProAsnCysIleLeuPheIleGluAsnLeuThrAlaGlnInTyrGl  577
     |||||||||||||||||||||||||||||||||||||||||||||||||
1821 AGCCCCAACTGCATCCTGTTCATCGAGAACCTCACGGCGCAGCAGTACGG 1870

 577 yHisTyrPheCysGluAlaGlnGluGlyTyrPheArgGluAlaGlnH  594
     |||||||||||||||||||||||||||||||||||||||||||||||||
1871 CCACTACTTCTGCGAGGCCCAGGAGGGCTCCTACTTCCGCGAGGCTCAGC 1920

 594 lsTrpGlnLeuLeuProGluAspGlyIleIleMetAlaGluHisLeuLeuGly  610
     |||||||||||||||||||||||||||||||||||||||||||||||||
1921 ACTGGCAGCTGCTGCCCGAGGACGGCACCATCATGGCCGAGCACCTGCTGGGT 1970

 611 HisAlaCysAlaLeuAlaAla 617
```

```
  11 SerAlaGlnGlyHisLeuArgSerGlyProArgIlePheAlaValTrpLy  27
     |||||||||||||||||||||||||||||||||||||||||||||||||
 171 TCCGCCCAGGGCCACCTAAGGAGCCGGACCCCGCATCTTCGCCGTCTGGAA 220

  27 sGlyHisValGlyGlnAspArgValAspPheGlyInThrGlnProHisT  44
     |||||||||||||||||||||||||||||||||||||||||||||||||
 221 AGGCCATGTAGGGCAGGACGGGGTGGACTTTGGCCAGACCGCCACA 270

  44 hrValLeuPheHisGluProGlyGlySerSerValTrpValGlyGlyArg  60
     |||||||||||||||||||||||||||||||||||||||||||||||||
 271 CGGTGCTTTTCCACGAGCCAGGCAGCTCCTCTGTGTGGGGAGGACGT 320

  61 GlyLysValTyrLeuPheAspPheProGluGlyLysAsnAlaSerValAr  77
     |||||||||||||||||||||||||||||||||||||||||||||||||
 321 GGCAAGGTCTACCTCTTTGACTTCCCCGAGGGCAAGAACGCATCGTGCG 370

  77 gThrValAsnIleGlySerThrLysGlySerCysLeuAspLysArgAspC  94
     |||||||||||||||||||||||||||||||||||||||||||||||||
 371 CACGGTGAATATCGGCTCCACAAAGGGGTCCTTGTCTGGATAAGCGGGACT 420

  94 ysGluAsnTyrIleThrLeuGlyArgArgSerGluGluValGlyLeuLeuAla 110
     |||||||||||||||||||||||||||||||||||||||||||||||||
 421 GCGAGAACTACATCACTCTCCTGGAGAGGCGGAGTGAGGGGCTGCTGGCC 470

 111 CysGlyThrAsnAlaArgHisProSerCysTrpAsnLeuValAsnGlyTh 127
     |||||||||||||||||||||||||||||||||||||||||||||||||
 471 TGTGGCCACCAACGCCCGGCACCCCAGCTGCTGGAACCTGGTGAATGCAC 520

 127 rValValProLeuGlyLysGluMetArgGlyTyrAlaProPheSerProAspG 144
     |||||||||||||||||||||||||||||||||||||||||||||||||
 521 TGTGGTGCCACTGGGCAGGATGAGAGGCTATGCCCCCTTCAGCCCGGACG 570

 144 luAsnSerLeuValLeuPheGluGluValThrGluValTyrSerThrIleArg 160
     |||||||||||||||||||||||||||||||||||||||||||||||||
 571 AGAACTCCCTGGTTCTGTTTGAAGGGGACAGGAGGGTGATTCCACCATCCGG 620

 161 LysGlnGluTyrAsnGlyLysLysIleProArgPheArgArgIleArgGlyGl 177
     |||||||||||||||||||||||||||||||||||||||||||||||||
 621 AAGCAGGAATACAATGGGAAGAAGATCCCGGTTCCGCCGCATCCGGGGCGA 670

 177 uSerGlyGluLeuTyrThrSerAspThrValMetGlnAsnProGlnPheIleL 194
     |||||||||||||||||||||||||||||||||||||||||||||||||
 671 GAGTGAGCTGTACACCAGTGATACTGTCATGCAGAACCCACAGTTCATCA 720

 194 ysAlaThrIleValHisGlnAsnProGlnAlaThryrAspSerPLysIleTyrTyr 210
     |||||||||||||||||||||||||||||||||||||||||||||||||
 721 AGGCCACCATCGTGCACCAAGACCCAGCTACGATGACAAGATCTACTAC 770

 211 PhePheArgGluAspAsnProAsnProLysAsnProGluAlaProLeuAsnVa 227
     |||||||||||||||||||||||||||||||||||||||||||||||||
 771 TTCTTCCGAGAGGACAATCCTGACAAGAATCCTGAGGCTCCTCTCAATGT 820

 227 lSerArgValAlaGlnAsnLeuCysArgGlyAspGlnGlyGlyGluSerSerL 244
     |||||||||||||||||||||||||||||||||||||||||||||||||
 821 GTCCCGTGTGGCCCAGTTGCAGGGGGACCAGGGTGGGGAAAGTTCAC 870

 244 euSerValSerLysTrpAsnThrPheLeuLysAlaMetLeuValCysSer 260
     |||||||||||||||||||||||||||||||||||||||||||||||||
 871 TGTCAGTGTCTCCAAGTGGAACACTTTTCTGAAAGCCATGCTGGTATGCAGT 920

 261 AspAlaAlaThrAsnLysAsnPheAsnArgLeuGlnAspValPheLeuLe 277
     |||||||||||||||||||||||||||||||||||||||||||||||||
 921 GATGCTGCCACCAACAAGAACTTCAACAGGCTGCAAGACGTCTTCCTGCT 970

 277 uProAspProSerGlyGlnTrpArgAspThrArgValTyrGlyValPheS 294
     |||||||||||||||||||||||||||||||||||||||||||||||||
 971 CCCTGACCCCAGCGGCCAGTGGAGGGACACCAGGGTCTATGGTGTTTTCT 1020

 294 erAsnProTrpAsnTyrSerAlaValCysValTyrSerLeuGlyAspIle 310
     |||||||||||||||||||||||||||||||||||||||||||||||||
1021 CCAACCCCTGGAACTACTCAGCCGTCTGTGTGTATTCCCTCGGTGACATT 1070
```

```
                     ||||||||||||||||||||
1971 CATGCCTGTGTCCCCTGGCTGCC 1991

seq_name: N_Geneseq_36:X02665

seq_documentation_block:
ID  X02665 standard; DNA; 8888 BP.
AC  X02665;
DT  07-MAY-1999 (first entry)
DE  EP-892047 Seq ID 41.
KW  Semaphorin L; human; immunosuppressant; anti-inflammatory; gene therapy;
KW  organ transplantation; inflammation therapy; immunotherapy; agonist;
KW  immunomodulatory; antagonist; ss.
OS  Homo sapiens.
PN  EP-892047-A2.
PD  20-JAN-1999.
PF  06-JUL-1998; 112470.
PR  11-FEB-1998; DE-005371.
PR  09-JUL-1997; DE-029211.
PA  (HMRI ) HOECHST MARION ROUSSEL DEUT GMBH.
PI  Ensser A, Fleckenstein B;
DR  WPI: 99-083564/08.
PT  New semaphorin L proteins - used as immunosuppressants and
PT  antiinflammatory agents in organ transplants, inflammation therapy,
PT  immunotherapy and gene therapy
PS  Disclosure; Page 106-111; 135pp; German.
CC  This invention describes a novel human semaphorin L protein. This protein
CC  or its encoding DNA are useful as immunosuppressants and/or
CC  anti-inflammatory agents in organ transplantation, inflammation therapy,
CC  immunotherapy and gene therapy. The DNA can be used to produce knock-out
CC  or knock-in animals for research purposes. The proteins or DNA can be
CC  used to search for the corresponding receptors or to screen for
CC  immunomodulatory agonists or antagonists.
SQ  Sequence  8888 BP;  1773 A;  2359 C;  2758 G;  1998 T;

alignment_scores:
      Quality: 1907.50          Length:    2757
      Ratio:   3.222            Gaps:        18
Percent Similarity: 21.473    Percent Identity: 21.364

alignment_block:
US-09-041-236-2 x X02665

Align seg 1/1  to: X02665  from: 1  to: 8888

  41 GluProHisThrValLeuPheHisGluProGlySerSerValTrpVa 57
     |||||||||||||||||||||||||||||||||||||||||||||||||
   1 GAGCCGCACCGGTGCTTTTCCACGAGCCAGGCCAGCTCCTCCTGTGTGGGT 50

  57 lGlyGlyArgGlyLysValTyrLeuPheAspPhePheGluProGluGlyLysAsnA 74
     ||||||||||||||||||||||||||||||||||||||||||||||||
  51 GGGAGGACGTGGCAAGGTCACCCTTTGACTTCCCGAGGGCAAGAACG 100

  74 laSerValArg 77
     |||||||||||
 101 CACTGTGCGCACGGTGAGCCTCTCTCTCCCCAACACCCCCCTACCC 150

  77 ..................................................... 77
 151 TCTTATCTCCCCCTGGCCCTGCCCAGGGTCCTCAGGGAATCCGAGGGAG 200

  77 ..................................................... 77
 201 CTGGCTTCTTCCTAAACTGCCCCCACCTCCGTATCCTATAAATGGCTC 250

  77 ..................................................... 77
 251 CTGGGGAGGAGGCTCCCTAAAGGTAGTCCAGATTGGAGTGGGGAGCTGGGGC 300

  77 ..................................................... 77
 301 GGTGTGGGAGAAAAACAGGAGCTAATGGGCCTGGCCAGCTGGCAGCGCTG 350

  77 ..................................................... 77
 351 CTGCCGGAAAGCCCAGGCTGGAAGCTGGGCCCCAGAGCCCATGCCTGGTCT 400

  77 ..................................................... 77
 401 TCTGAACCCTCTGGGCCTCAGCTCTGGATATGAGACCCTGTTTGACCTCA 450

  77 ..................................................... 77
 451 GGTAGATCACTCACCCTCTCAGAGCCCCAGTTGCTCATCGTCAGATGAG 500

  77 ..................................................... 77
 501 AATAATCGTTGCTTCCTTTGGGGCTTATCCTGAGGCTGTGTGGAAAGCAT 550

  77 ..................................................... 77
 551 TYCAGGGTACCTCACCCTGCAGATTGAACTAATGCTTCTCCCCTCC 600

  78 Thr.ValAsnIleGlySerThrLysGlySerCysLeuAspLysArg... 92
     ||||||||||||||||||||||||||||||||
 601 CCAGGTGAAATCGGCTCCACAAAGGGTCCTGTCTGGATAAGCGGGTGA 650

  92 ..................................................... 92
 651 GCGGGGGAGGGGATCTGGAGGGGTCTGAGCCCACTTGGTAAAGGGAGAAGGAG 700

  92 ..................................................... 92
 701 ACCCTGAGGGTCTAAGGAAGGAAGCATGGCCCCTGCCCCACGAGTCCCAGA 750

  92 ..................................................... 92
 751 CTGATGGGGGAGGACGTGTCCTCGTGCTTAGGGGATGGCCGTCAGCTGCAC 800

  92 ..................................................... 92
 801 ACACTCTGGCGTGTCCCGGGAGGCGTGCACCTATGCTAAGCCCTTCTGAC 850

  92 ..................................................... 92
 851 ACCTTCTTCCCCTGATCCTGGGGGTCCTCAGTGCTGCTAGGCTTGCCAGGGCCTT 900

  93 ...........................................AspCysG 95
                                               ||||||
 901 CCAGCAACCAATTTCTCTCCTCCCCTTCTCCTCGCGGGCAGGAGACTGCCG 950

  95 luAsnTyrIleThrLeuLeuGluArgSerGluGlyLeuLeuAlaCys 111
     |||||||||||||||||||||||||||||||||||||||||||||||
 951 AGAACTACATCACTCTCCTGGAGAGGCGAGTGAGGGGCTGCTGGCCTGT 1000

 112 GlyThrAsnAlaArgHisProSerCysTrpAsnLeu.......... 123
     |||||||||||||||||||||
1001 GGCACCAACGCCCGGCACCCCAGCCTGCTGGAACCTGGTGAGAGGCTGCT 1050

 123 ..................................................... 123
1051 CCCCATGTGCCTGATCAGCTCACCTTCTACTGCGTGGGCTTCTGCCCCTC 1100

 123 ..................................................... 123
1101 ATGGYGGGAAGGAGGATGGCGAGACTCCAATGCTGCCCTGGCCTGGGAGG 1150

 123 ..................................................... 123
1151 ATGGGGGCTCCTGGCCGAGAAACTGCCCTCATGGGAGGCAGTGGCTGTGG 1200

 124 ..............................ValAsnGlyThrV 128
                                   |||||||||
1201 GATTATGTGGCCATCCAACCCTCTGGATCTCCCACAGTGGAATGGCACTG 1250
```

```
128 alValProLeuGlyGluMetArgGlyTyrAlaProPheSerProAspGlu 144                                              2151 CCCATTCGCTTCTGTGCCCTGGCGACCCACAGTTCATCAAAGCCACCAT 2200
    |||||||||||||||||||||||||||||||||||||||||||||||||                                                       |||    ||| ||||||||||||||||||||
1251 TGGTGCCACTTGGCGAGATGAGAGGCTACGCCCCCTTCAGCCCGGACCAG 1300                                               197 eValHisGlnAspGlnAlaTyrAspAspLysIleTyrTyrPhePheArg 214
                                                                                                             |||||||||||||||||||||||||||||||||||||||||||||||||
145 AsnSerLeuValPheGlu................................. 151                                               2201 CGTGCACCAAGACCAGGCTTACGATGACAAGATCTACTACTTCTTCCGAG 2250
    ||||||||||||
1301 AACTCCCTGGTTCTGTTTGAAGGTGGGCATGCTTCGGAACTGGGCTGG 1350                                                214 luAspAsnProAspLysSerAspGluAlaProLeuAsnValSerArgVal 230
                                                                                                             |||||||| || | ||||||||||||||||||||||||||||||||||
151 ................................................. 151                                               2251 AGGACAATCCTGACAAGATCCTGAGGCTCCTCTCAATGTGTCCCGTGTG 2300
1351 GAGCAGGATGGTCAGCTCTTTGTCCAGTGTCCGGAGGAGGGACTTCCAGG 1400                                                231 AlaGlnLeuCysArg................................... 235
                                                                                                             |||||||||||||
152 ...........................GlyAspGlu 154                                                             2301 GCCCAGTTGCAGGGTGAACACGGGGTGGAGGGCTGCTGGCTACGTGTC 2350
                               ||| ||||
1401 AGCTGCCTGCCCTTACTCATTTCTCCCTCCCACTGACCCCAGGGGACGAG 1450                                               235 ................................................. 235
1451 GTGTATTCCACCATCCGGAAGCAGCAGGAATACAATGGGAAGACATCCCTCCGGTT 2351 TGTGCATGAATAGGCCTGAGTGAGGGTGAGTTCTGTGTGTCCGTGTGCAT 2400
155 ValTyrSerThrIleArgLysGlnGluTyrAsnGlyLysIleProArgPh 171
    |||||||||||||||||||||||||||||||||||||||||||||||||                                                    235 ................................................. 235
171 eArgArgIleArgGlyLeuSerGluLeuTyrThrSerAspThrValMetG 188                                                2401 GTAGAAGTGTGTGGATGTATGAGTGAGTGGTCTGTGTCAGGGACTGTGGGAG 2450
    |||||||||||||||||||||||||||||||||||||||||||||||||
1501 CCGCCGCATCCGGGGCCGAGAGTTGAGCTGTGATACTGTCATGC 1550                                                    235 ................................................. 235
188 lnAsn............................................. 189                                                2451 CAGCTGTGTGTCGCATGAGGAGCATCATGTCTGTCTGTGTGGTAAAGGTGGC 2500
    || :::
1551 AGAGTGAGTCAGGCTCCGGCTGGGCTGAGGGTGGGCAAGGGGTGTGAGC 1600                                                235 ................................................. 235
189 ................................................. 189                                                2501 TGAGCTCCTGCACGCTATGAGCGTGTGAGCGTGTGTGATGGGGTG 2550
1601 ACTTAAGGTGGCAGATGGGATCCTGATGTTCTGGGACGGCCTCCCTGAGG 1650                                                235 ................................................. 235
189 ................................................. 189                                                2551 TGTGTGTGTGTGTGTGTGTGTTTGCCTGTGTGTGAATGTGCTGTGCCAC 2600
1651 GCCGCTGGGGCCATGCAGCAGGAAAGCAGGACCCTTGGTATAGGCCTGAGAAGT 1700                                             235 ................................................. 235
189 ................................................. 189                                                2601 GTATGTGGGTGCGTGAGTCAGTAAATGTGTCTCAGTTCCGTCCGTCTGCTCTG 2650
1701 TAGGGTTGGCTGGGAGCAGAGAACAGACAAGGTATAGCAGTGGGATGGG 1750                                                235 ................................................. 235
189 ................................................. 189                                                2651 TGGGGACCTGGCACTCTCACCTGCCCTGGGCCACTGCTGGCCCCTG 2700
1751 CCCAGCCCCTCTTCAGGAACACAAACAGAGGGAGCCCCAGTGCCAGTGCAG 1800                                               235 ................................................. 235
189 ................................................. 189                                                2701 GGCTCTGGATCAGCCAGGCCTGCTTGCAGGAGTCTCATCTGGAGACCTGC 2750
1801 GGTCCCCAGGAGCCAAAGTTTATCCTCTGCTGAGTTCACGTGGAGGCAGC 1850                                                235 ................................................. 235
189 ................................................. 189                                                2751 CCTGAGTCCTGGGGCACCCCCGGCAGGTCCTGCCCCCTCGCCAGCCTGCCT 2800
1851 CCCCCAACTCCCTCCTCATCAGGGCTCTGCCAATTGAGCAGAAGTGACAT 1900                                               235 ................................................. 235
189 ................................................. 189                                                2801 TCCTCCCTCTGGGCCCCAGTGTTGATATTGCTGGCAGTGGTTTCCTGGGGT 2850
1901 AGGGGGCCCCCAGGGGACCTTCCCCCACTCCCCAGGCCATGAAGTCATTGCTC 1950                                            235 ................................................. 235
189 ................................................. 189                                                2851 GTGTGGGGAAGCCCGGCAGGGCTGAGGGGCCCTTCTTCCCCCTCTCCACCC 2900
1951 CTGGGCCGATGACATCTTTGTAGGAAGAGGGCAAAACAGGTGTGGGGTGG 2000                                                236 ..........GlyAspGlnGlyGlyGluSerLeuSerValSerLysTrpAs 250
                                                                                                             ||||||||||||||||||||||||||||||||||||||||
189 ................................................. 189                                                2901 TTCCAGGGGACCAGGCGGTGGGGGAAAGTTCACTGCAGTTCCAAGTGGAA 2950
2001 AGGTGCAGGGTCTAGGGCCCCTCGGGAGTTGGACCTGATGTTATGAGTC 2050
                                                                                                            250 nThrPheLeuLysAlaMetLeuValCysSerAspAlaAlaThrAsnLysA 267
189 ................................................. 189                                                     |||||||||||||||||||||||||||||||||||||||||||||||||
2051 CTATTCCAGATCTGATTTGCCATGGTTTGTGCAGACCCGAAGGAGGGAGG 2100                                                2951 CACTTTTCTGAAAGCCATGCTGTATGCAGTGATGCTGCCACCAACAAGA 3000
189 ................................................. 189                                                267 snPheAsnArgLeuGlnAspValPheLeuLeuProSerProSerGlyGln 283
                                                                                                             |||||||||||||||||||||||||||||||||||||||||||||||||
2101 AGAGTGTGCAGGGTTGGAAATGTGCTCCCCGGGCAAGCTTCCCAGCCCTTACG 2150                                             3001 ACTTCAACAGGCTGCAAGACGTCTTCCTTCTGCCCTCCCCGACCCCAGCCGGCCAG 3050
190 ...........Pro.GlnPheIleLysAlaThrIl 197                                                              284 TrpArgAspThrArgValTyrGlyValPheSerAsnProTrp....... 297
                                                                                                             |||||||||||||||||||||||||||||||
```

```
316  hrSerSerLeuLysGlyTyrHisSerLeuProAsnProArgProGly 332
     |||||||||||||||||||||||||||||||||||||||||||||||
4001 CCTCCCTCACTCAAGGGCTACCACTCAAGCCCTTCCCAACCCGCGGCCCTGGC 4050

332  ................................. 332
4051 AAGGTGAGCGTGACACCAGCCGTGGCCCAGGCCCTCCTTCTGCCT 4100

333  ...Lys..CysLe 335
         ||| ||
4101 CACCTCCCACCACCCCACTGACCTGGGCCTGCTCTCCTTGCCCAGTGCCT 4150

335  uProAspGlnGlnProIleProThrGluThrPheGlnValAlaAspArgH 352
     |||||||||||||||||||||||||||||||||||||||||||||||
4151 CCCAGACCAGCCAGCCCGATACCCACAGAGACCTTCCAGTGGCTGACCGTC 4200

352  isProGluValAlaGlnArgValGlnProMetGlyProLeuLysThrPro 368
     |||||||||||||||||||||||||||||||||||||||||||||||
4201 ACCCAGAGGTGGCGCAGAGGGTGGAGCCCATGGGGCCTCTGAAGACGCCA 4250

369  LeuPheHisSerLysTyrHisTyrGlnLysValAlaValHisArgMetGl 385
     ||||||||||||||||||||||||||||||||||||||||||||||||
4251 TTGTTCCACTCTAAATACCACTACCAGAAAGTGGCCGTCCACCGCATGCA 4300

385  nAlaSerHisGlyGluHisPheHisValLeuTyrLeuThr.......... 398
     |||||||||||||||||||||||||||||||||||
4301 AGCCAGCCACGGGGAGCACTTTCATGTGCTTTACCTAACTACAGGTGAGA 4350

398  .................................... 398
4351 GGCTACCCCGGGACCCTCAGTTGCTTTGTAAAAACGGCATGAAAGGTG 4400

398  .................................... 398
4401 TAAGGAATAATGTAGTTAACATCTGGTTGGATCTTTACATGTGGAAGGAA 4450

398  .................................... 398
4451 TAATTGAGTGACTGAGTTGTCAGGGGTTAATGTGTGTGGGTGTGGAAGA 4500

398  .................................... 398
4501 GCCAGGCAGGGAGAGCTTCCTGGAGGAGGTAGGGGCAAGAGGAAAGGGG 4550

398  .................................... 398
4551 GATGGGAGAAAAGCAGCACTGGGATTTGGAGGCCGAAATCTGGAGAGTC 4600

398  .................................... 398
4601 TGAGCAAAGCCAGGTGCACCTTTGGTCCAGATGTCTGACTCAGGGAAGAA 4650

398  .................................... 398
4651 GATGGTAGGAGAGACGTGGCAAATGAGGAGGAGGGGCCTGAACCACACAGG 4700

398  .................................... 398
4701 GATACTGGCCTCTGCCAGGCAGAATGAGGAGTCAGGCCCTGCGCCTGTC 4750

398  .................................... 398
4751 TTTGGGATTGTGCAGGTGAGAAGAAACATTTGAGGAGTTGATGGGCACA 4800

398  .................................... 398
4801 AATTAGGTATGGGAAGGAGTTCCAGGGGGCAGAACCTTTGCCATCTCAC 4850

398  .................................... 398
4851 AGAGGACAGGGCAGCTTCTCTTCTTCCCTGGAGTAGGCCCTGCTGGGGG 4900
```

```
3051 TGGAGGACACCAGGGTCTATGGTGTTTCTCCAACCCCTGGTGAGTGGC 3100
297  ................ 297
3101 CCTTGTCCTGGGGCCGGGGCTGGCATTGGTTCAGTGTCCAGTAGGGACAG 3150
297  ................ 297
3151 GAGGGCCTTGGGCCCTGCTGAGGGGCCTCCCTGGTGTGGCAGGAGCAGGGGC 3200
297  ................ 297
3201 TGCAGGCTCAAGAGAGCTGGGCTGTGCTGCTGGGTGTGGGGGGGACAGC 3250
297  ................ 297
3251 CAGTGCGATGTATGTACTGTTGTGTGAGTGAGTCTGCACTCATGGGTGTG 3300
297  ................ 297
3301 TGTGCATGCCCTATATGCACACTCATGACTGCACTTGTGCCTGTGTGTCC 3350
297  ................ 297
3351 CACACCCTGCTGTTGTGCCGAGAGTGACACTGGGCCCAGGAGGAAGCTGCT 3400
297  ................ 297
3401 GAAGCATCTCTCGGGGAGCTGGGTGCTATTACACCTGCTCAGGCACTGCC 3450
297  ................ 297
3451 TGAGCCCGATAATTCACACTTCTTAATCACTCTCATTGATTGAACACACG 3500
297  ................ 297
3501 GCAGGCGGAAGTGTTGGGGTGTGTGGGGAGGAGTTAGGGATAGAGTGGAG 3550
297  ................ 297
3551 GAAGCCAAGAGACCCTGCTGCTCTGGGTCCTGGGTGGGTCCCCCAGGCT 3600
297  ................ 297
3601 GGGAAGGGGTTGGGGGTCTGCCCTCTGCCCTCCTGGGCATCAGCAGGCCT 3650
297  ................ 297
3651 GTGCCCAGGGAGGGGCTAGAGAACTGCTCAGCCTATGATGGGGTTCCTCCT 3700
297  ................ 297
3701 GCCTTGGGGTTGGGTAGAGCAGATGGCCTCTCAGTGATTCTGTAA 3750
297  ................ 297
3751 CAGGATACAAGTTTGTGGTTTAAATTGCAGCACAAAGAAATTAGGCTGA 3800
297  ................ 297
3801 ACTCCTCTCCTTCCTCCCTCCCATCCCTCCCCATTTCAGTGGTGGTTGG 3850
297  ................ 297
3851 CAACTCAGTGCCAGGCACAAGGCTGGCCTGGGTGAGGTGGAGGTGGATGGG 3900
298  ...Asn.Ty 299
        :: ||
3901 TGGGTTCTGGGGCCCCCATTGAGCTGGTCTCCATGTCACTGCAGGAACTA 3950
299  rSerAlaValCysValTyrSerLeuGlyAspIleAspLysValPheArgT 316
     |||||||||||||||||||||||||||||||||||||||||||
3951 CTCAGCCCGTCGTGTGTATTCCCTGGGTGACATTGACAAGGTCTTCCGTA 4000
```

```
    398  ..................................................
   4901  AAGCTGGGTGGAATGCCGTGGAGATGCTCCTGCTTTCTGGGAAAGCCACA  4950
    398  ..................................................
   4951  GGACACGGAGGAGCCAGTCCTGAGTTGGTTGTCGGCAGCTTCCCCATGCC  5000
    398  ..................................................
   5001  AGCTGCCTTCCTTGAGACTGGAAAGGGCCTCTAGCACCCCTGGGGCCATT  5050
    398  ..................................................
   5051  CAATTCAGGCCCAGGCGCCCAACCTCAGTGTTCACATTCCCATGTGAT   5100
    398  ..................................................
   5101  CTCCTGTTGCTGCTTCACCTTGGGACTGTCTCGGCTTTGGTGACCTTGTA  5150
    398  ..................................................
   5151  GGAAACTGGAACCCCAGCACCATTGTTTGGGCTCCTGGAAGCCTTGGGGAG  5200
    398  ..................................................
   5201  AGGAATTCCCACAGGGCCAGGGCCTGGGTCCTGATTCCCTGCCTCTTTAC  5250
    398  ..................................................
   5251  TCCCTATTCATCCGGCTACACCCTTGGGCCCCCATCCTTGCTGGCTCC   5300
    398  ..................................................
   5301  AGTACTGCTGGCACAGCTGTTGTGGTCATCCAGGGATGGCAGGGGCACTG  5350
    398  ..................................................
   5351  GGGAACAGAAGAGAGAGGTCACACAGTGCGAACTGGGAGCAGGAGCTAG   5400
    398  ..................................................
   5401  GACAAGGAAGGCTGGACTTGGGCCCATGATTCCCTTCCTGCAGACTTGGG  5450
    398  ..................................................
   5451  AAGTGAGGACACTTGAGTGATTAGAGAGAAGGTGTCTTCGTTCTAAGGGCAG  5500
    398  ..................................................
   5501  TGGAGGAGGCACCATTTTGGGAGCCTGCATCATTCGTATTTGGGCTAGATT  5550
    398  ..................................................
   5551  GAAAAATAGAGCTTTCTAAGTCCTCGCAGAGAATGGGAGGCTCTCACAA   5600
    398  ..................................................
   5601  CTGGGAGAAGTATTGGCTCTTTTCCTGAGAATTTTGCCAAGGGTATGCTG  5650
    398  ..................................................
   5651  TTACTGGGGCTGGTTTGGAAGGAGTATAGGGCCATTATGTCTGTGAAGGCA  5700
    398  ..................................................
   5701  GTGGCTGGGGTTGGGGTGGGGCCTTATCAGGGCCCAAGAGGAGCATCTCTC  5750
    398  ..................................................
   5751  AGAGTCCACAGATGAGGATCACGGATGTGTGAGGAAACATCCTAGGCAG   5800
    398  ..................................................

   5801  GCAATCATCTGACTGCTGCTTTTTGGGGCAGGTGATGCCCTGGGAAATTGGG  5850
    398  ..................................................
   5851  AGGGAGGGAGAGAGGGAGGTAGGCTATTCTAGAAACTGGGAGAGCAGGTG  5900
    398  ..................................................
   5901  AGTGTAGGATTGGGAGGACCAGGGTGTCAGGGTCCCCATTGGTCCCTAATTG  5950
    398  ..................................................
   5951  AGAACGGAGAGAGAGCGCATTGGTCTAGGAGGCAGGCAGCTCGGTTATAAGACC  6000
    398  ..................................................
   6001  TTGGGAACTCCTTGATTTAGAATCCAAGATCCTTTTTAGATCTAGGATTTT  6050
    398  ..................................................
   6051  ATAAAATTAAGATATCCCCTAAGATCAAATGCAACGTGAGTCCTGAATT   6100
    398  ..................................................
   6101  GGATCCTAGAACACAGAAGAGGACATTTGTGGAAAAACTAGTGAAATCCAA  6150
    398  ..................................................
   6151  ATAAAGTCTGTAGTTTGTTAATAGTAATGCACCAATGTCAGTTGCCTAG   6200
    398  ..................................................
   6201  TTGTGACAAATATACCGTGGTTATGTGAAGATGGTAAACATTAGGGGGAACT  6250
    398  ..................................................
   6251  GGAGAAGGGTAGATTGGAGCTCCTCGTACTACATCTTTGCAACTTTTCTGGG  6300
    398  ..................................................
   6301  AATCTAAAATTACTCCAAAATAAAAAAAATGTATTTAAAGTAAATATA    6350
    398  ..................................................
   6351  TTCCCTAAGAGTCCAGGAGGCAGGGGACGTTGTAGAAGCAGCTGAGTGGTT  6400
    398  ..................................................
   6401  GGGTTCTGACAGATTTGGTTCCAACTCGGTCTCTGCTGCTCACCAGCTGT  6450
    398  ..................................................
   6451  GTGACCTTGAGCAAGTGGCTTAGCCTTTCGAGCCTGATTCCTTATCTG   6500
    398  ..................................................
   6501  TGGAGTGGGGAAGATGAGCAGCACCTCGCAGGGCGTGTGGAGGGTTAAACG  6550
    398  ..................................................
   6551  AGTGATGATGCATGGACAGCAGCCGCCACTGACCTTGCGTGGTGTGGGGCTCCT  6600
    398  ..................................................
   6601  GCTTCTGTTCTTCCCGTGCAGCCTTGGGAATGTTGGAGGCCGTATCCAGG  6650
    398  ..................................................
   6651  GACCCCTGGGCCTTCTCTGGATGCCTTCTCTGATCAGCCTTGGAAGGTTC  6700
    398  ..................................................
```

```
6701 CAGGCTGCCCTAGGCTCCCACATTCTTCCCCAGTCACGCTCTCCTCGCC 6750
 398 .................
6751 CTGCCACACCAGTCCTGTGACCCTTGCCTGAGTTGTGACTTCCCACCCC 6800
 398 .................
6801 TCCCCGGCCTAGAGGAAAGCTGCCTGGCCCCTCAGTGGGACTCCCGCCCA 6850
 399 .......Thr.AspArgGlyThrIleHisValVa 408
               |||  ||||||||||||||||||
6851 CTGACCCTCGTCCACCATACACAGCACGGGCACTATCCACAAGGTGGT 6900
 408 lGluProGlyGluGlnHisSerPheAlaPheAsnIleMetGluIleG 425
     |||||||||||||||||||||||||||||||||||||||||||||||
6901 GGAACCGGGGAGGAGCGAGGCACACAGCTTCGCCTTCAACATCATGGAGATCC 6950
 425 lnProPheArgAlaAlaAlaAlaIleGlnThrMetSerLeuAspAlaGlu 441
     ||||||||||||||||||||||||||||||||||||||||||||||||||
6951 AGCCCTTCCGCCCGCGCCGGCTGCCATCCAGCACCATGTCGCTGGATGCTGAG 7000
 442 Arg.............. 442
     |||
7001 CGGGTGAGCCTTCCCCCACTGCGTCCCATGGGCTATGCAGTGACTGCAGC 7050
 442 .................
7051 TGAGGACAGGGCGTCCTTGCATGTGATTTGTGTGTCTTTAAGAGCTTC 7100
 442 .................
7101 TAGGCCTTAGGGCCTGGACATTTAGGACTGAGTGTGGGGTGGGGCCCGGG 7150
 443 .......ArgLysLeuTyrValSerSerGl 450
            ||||||||||||||||||||||
7151 CCTGACCCAATCCTGCTGCCTTCCAGAGGAAGCTGTATGTGAGCTCCCA 7200
 450 nTrpGluValSerGlnValProLeuAspCysLeuCysValTyrGlyGlyG 467
     |||||||||||||||||||||||||||||||||||||||||||||||||
7201 GTGGGAGGTGAGCCAGGTGCCCCTGTGTGAGGTCTATGCGCGGGG 7250
 467 lyCysHisGlyCysLeuMetSerArgAspProTyrCysGlyTrpAspGln 483
     ||||||||||||||||||||||||||||||||||||||||||||||||||
7251 GCTGCCACGGTTGCCCTCATGTCCCGAGACCCCTACTGCGGCTGGGACCAG 7300
 483 .................
7301 GGGCGCTGCCATCCATCCATTCCTTCCAGAGCTCCGAACGGTACGTTGGCCCGGGATCC 7350
 484 GlyArgCysIleSerIleTyrSerSerGluArg........... 494
     |||||||||||||||||||||||||||||||||
7351 CTCCGTCCCTGGGACAAGGTGGGCATGGGACATGGGGAGGTGTGTCGGG 7400
 494 .................
7401 CTGGAAGAGGTGGCGGTACTGGGCCTTTCTTGTGGGACCTCCTCCTCTACT 7450
 494 .................
7451 GGAACTGCACTAGGGGTAAGGATAATGAGGGTCAGGTCGCAGCCTTGTAT 7500
 495 .........Ser..ValLeuGlnSerI 500
                ::: ||||||||||||
7501 CTGCTGATCCTCTTTCGTCCTTCCCACTCCAGGTCGCAATCCAT 7550
 500 eAsnProAlaGluProHisLysGluCysProAsnProLys.......... 513
     ||||||||||||||||||||||||||||||||||||||||
7551 TAATCCAGCCGAGCCACCACAAGGAGTGTCCCAACCCCAAACCAGGTACCT 7600
 513 .................
7601 GATCTGGGCCTGCTGCGGCGTGTGCGGCCCAATGAGTGGGGTACTGCCCTGC 7650
```

```
 513 .................
7651 CCTGATTGTCCTGGTCTGAGGGAAACATGGCCTTGCCTGCTGGGGCCCCAG 7700
 513 .................
7701 GTACATGGGCAGGATACAGTCCTGCAGAGGGAGCCCTCTTGGTGGGATG 7750
 513 .................
7751 AGCGAGACGGGAGAAAAAGGAGGACGCTGAGGGGCTGGGTTCCCCACGTT 7800
 513 .................
7801 CATTCAGAAGCCTTGTCCTGGGATCCCAGTCGGTGGGGAGGACACATCCT 7850
 514 ...............P 514
7851 CCCCTGGGAGCTCTTTGTCCCTCCTCACGGCTGCCTTCCCCACTGCCTCCC 7900
 514 roAspLysAlaProLeuGlnLysValSerLeuAlaProAsnSerArgTyr 530
     ||||||||||||||||||||||||||||||||||||||||||||||||||
7901 CAGACAAGGCCCCACTCGCAGAAGGTTTCCCTGGCCCCCAAACTCTCGCTAC 7950
 531 TyrLeuSerCysProMetGluSerArgHisAlaThrTyrSerTrpArgH 547
     |||||||||||||||||||||||||||||||||||||||||||||||||
7951 TACCTGAGCTGCCCCATGGAATCCCGCCACGCCACCTACTCCATGGCGCCA 8000
 547 sLysGluAsnValGluGlnSerCysGluProGlyHisGlnSerProAsnC 564
     |||||||||||||||||||||||||||||||||||||||||||||||||
8001 CAAGGAGAACCTGGACACAGACGCTGCGAACCTGGTCACCAGAGCCCCAACT 8050
 564 ysIleLeuPheIleGluAsnLeuThrAlaGlnGlnTyrGlyHisTyrPhe 580
     ||||||||||||||||||||||||||||||||||||||||||||||||||
8051 GCATCCTGTTCATCGAGAACCTCACGGCGCAGCAGTACGGCCACTACTTC 8100
 581 CysGluAlaGlnGluGlyTyrPheArgGluAlaGlnHisTrpGlnLe 597
     |||||||||||||||||||||||||||||||||||||||||||||||
8101 TGCGAGGCCCAGGAGGGCTCCTACTTCCGGGAGGCTCAGCACTGGCAGCT 8150
 597 uLeuProGluAspGlyIleMetAlaGluHisLeuLeuGlyHisAlaCysA 614
     |||||||||||||||||||||||||||||||||||||||||||||||||
8151 GCTGCCCGAGGACGGGCATCATGGCCGAGCACCTGCTGGGTCATGCCTGTG 8200
 614 laLeuAlaAlaAlaSerLeuTrpLeuGlyValLeuProThrLeuGly 630
     |||||||||||||||||||||||||||||||||||||||||||||||
8201 CCCTGGCCGCGCTCCCTCCTGGCTGGGGTGCTGCCACACTCACTCTTGGC 8250
 631 LeuLeuValHis 634
     ||||||||||||
8251 TTGCTGGTCCAC 8262
```

seq_name: N_Geneseq_36:X02628

```
PT  antiinflammatory agents in organ transplants, inflammation therapy,
PT  immunotherapy and gene therapy
PS  Disclosure: Page 60-61; 135pp; German.
CC  This invention describes a novel human semaphorin L protein. This protein
CC  or its encoding DNA are useful as immunosuppressants and/or
CC  anti-inflammatory agents in organ transplantation, inflammation therapy,
CC  immunotherapy and gene therapy. The DNA can be used to produce knock-out
CC  or knock-in animals for research purposes. The proteins or DNA can be
CC  used to search for the corresponding receptors or to screen for
CC  immunomodulatory agonists or antagonists.
SQ  Sequence 1195 BP;  254 A;  358 C;  336 G;  247 T;
```

```
alignment_scores:
     Quality: 1755.00              Length:        365
     Ratio:      5.072             Gaps:            2
Percent Similarity: 94.795     Percent Identity: 90.411
```

```
alignment_block:
US-09-041-236-2 x X02628   ..

Align seg 1/1  to: X02628  from: 1  to: 1195
```

```
   1 LeuLeuLeuLeuTrpAlaAlaAlaAlaAlaSerAlaGlnGlyHisLeuAr  17
     |||||||||::::|::||::||||||||||||||||||||| ||
 109 CTTCTGCTGGTGTTCTGGGTGGCCGCCGCCCTCCGCCCAAGGCCACTCGAG 158

  17 gSerGlyProArgIlePheAlaValTrpLysGlyHisValGlyGlnAspA  34
     ||||||||||||||||||:::::::||||||  ||||||||||
 159 GAGCGGACCCCGCATCTCCGCGTCTGGAAA.......GGGCAGGACC    199

  34 rgValAspPheGlyGlnThrGlyHisProHisThrValLeuPheHisGluPro 50
     ::|||||||:::|::||||||||:||||||||||||||||
 200 ATGTGGGACTTTAGCCAGGCCTGAGCCACACACCGTGCTTTTCATGAGCCG 249

  51 GlySerSerSerValTrpValGlyLysArgGlyLysLysValTyrLeuPheAs 67
     |||||  ||:::|||||||||||||||||||  |||::
 250 GGCAGCTTCCTTCGTCTGGGTGGTGGACGTGGCAAGGTCTACCACTTCAA 299

  67 pPheProGluGluProPheSerProAspGluAsnSerLeuValLeuSerT  84
     :|||||::|||||||||||||||||||||||||:|||::
 300 CTTCCCGGAGGGCAAGAATGCCTCTGCGCACGGTGAACATCGGCTCCA    349

  84 hrLysGlySerCysLeuAspLysArgAspCysGluAsnTyrIleThrLeu 100
     |||||||||||||::|::||||||||||||||||||
 350 CAAAGGGGTCCTGTCAGGACAACAGACTGTGGGAATTACATCACTCTT   399

 101 LeuGluArgArgSerGluGlyLeuGluLeuAlaCysGlyThrAsnAlaArgHi 117
     ||||||||||::|::::||:||||||||||||||||||
 400 CTAGAAAGGCCGGGGTAATGGGCGCTGGTCTGTGGCACCAATGCCCGGAA 449

 117 sProSerCysTrpAsnLeuValAsnGlyThrValVal...ProLeuGlyG 133
     |||||  ||:::||||||||||||||  |||||||
 450 GCCCAGCTGCTGGAACTTGGTGAATGACAGTGTGGTGATGTCACTTGGTG 499

 133 luMetArgGlyTyrAlaProPheSerProAspGluAsnSerLeuValLeu 149
     ||||::||::||||||||||||||||||||||
 500 AGATGAGAAGGCTATGCCCCCCTTCAGCCCGGATGAGAACTCCCTGGTTCTG 549

 150 PheGluGlyAspGluValTyrSerThrIleArgLysGlnGluTyrAsnGl 166
     |||||||||||||||::|:||||||||||||||
 550 TTTGAAGGAGATGAAGTGTACTCTACCATCCGGAAGCAGGAATACAACGG 599

 166 yLysIleProArgPheArgArgIleArgGlyGlyGluSerGluTyrThrS 183
     ||||||||||||||||||||||||||||||
 600 GAAGATCCCTCGGTTTCGACGCATTCGGGGCGAGAGTGAACTGTACACAA 649

 183 erAspThrValMetGlnAsnProGlnPheLeuLysAlaThrIleValHis 199
     ||||||||||||||||||||||||||||
 650 GTGATACAGTCATGCAGAACCCACAGTTCATCAAGGCCACCATTGTGCAC 699

 200 GlnAspGlnAlaTyrAspAspLysLysIleTyrTyrPheArgGluAspAs 216
     |||||||||||||||||||||||||||
```

```
 700 CAAGACCAAGCCTATGATGATAAGATCTACTACTTCTTCCGAGAAGACAA 749
     ||||||||||||||||||||||||||||||||||||||||||
 216 nProAspLysAsnProGluAlaProLeuAsnValSerArgValAlaAlaGlnLnL 233
 750 CCCTGACAAGAACCCCGAGGCTCCTCAATGTGTCCCAGTAGCCCAGT    799

 233 euCysArgGlyAspGlnGlyValGlyGlySerLeuSerValSerLysTrp 249
 800 TGTGCAGGGGGGACCAGGGTGTGAGAGTTCGTTCGTTCTCCAAGTGG    849

 250 AsnThrPheLeuLysAlaMetLeuValCysSerAspAlaAlaThrAsnLy 266
     ||||||||||||||||||||||||::::
 850 AACACCTTCCTGAAAGCCATGTTGGTCTGCAGCGATGCAGCCACCAACAG 899

 266 sAsnPheAsnArgLeuGlnAspValPheLeuLeuProAspSerGlyGlyG 283
     :||||||||||||||||||||||||||||
 900 GAACTTCAATCGGCTGCAAGATGTCTTCCTGCTCCCTGACCCCAGTGCC  949

 283 lnTrpArgAspThrArgValTyrGlyValPheSerAsnProTrpAsnTyr 299
     |||||||||||||||||||||||||||
 950 AGTGGAGAGATACCAGGGTCTATGGCGTTTCTCCAACCCCTGGAACTAC  999

 300 SerAlaValCysTyrSerLeuGlyLeuGlyAspIleAspLysProSerGlyArgTh 316
     ::::|:::::||:||||::::||||||||||::::
1000 TCAGCTGTCTGCGTGATTCGCTTGGTGACATTGACAGAGTCTTCCGTAC 1049

 316 rSerSerLeuLysGlyTyrHisSerSerLeuProAsnProArgProGlyLys 333
     |||||||||||||  :::||  |||||||||
1050 CTCATCGCTCAAAGGCTACCACATGGCGCTTTCCAACCCTGACCTGGCA  1099

 333 ysCysLeuProAspGlnGlnProIleProThrGluThrPheGlnValAla 349
     ||||||||||||||||||||||||||||
1100 TGTGCCTCCCAAAAAAGCAGCCCATACCCACAGAAACCTTCCAGGTAGCT 1149

 350 AspArgHisProGluValAlaGlnAsnArgValGluProMetGlyPro    364
     |||  |||||||||||||||||||||||||||
1150 GATAGTCACCCAGAGGGTGGCTCAGAGGGTGGAACCTATGGGGCCC     1194
```

```
seq_name: N_Geneseq_36:X02663
```

```
alignment_scores:
     Quality: 1069.00              Length:        200
```

Percent Similarity: 100.000   Percent Identity: 99.500

Ratio: 5.345   Gaps: 0

alignment_block:
US-09-041-236-2 x X02663

Align seg 1/1   to: X02663   from: 1   to: 4019

```
146  SerLeuValLeuPheGluGlyAspGluValTyrSerThrIleArgLysGl  162
     ||||||||||||||||||||||||||||||||||||||||||||||||||
148  TCCCTGGTTCTGTTGGAGGGACGAGGTGTATTCCACCATCCGGAAGCA     197

162  nGluTyrAsnGlyLysIleProArgPheArgArgIleArgGlyGluSerG  179
     ||||||||||||||||||||||||||||||||||||||||||||||||||
198  GGAATACAATGGGAAGATCCCTCGGTTCCGCCGCATCCGGGGCGAGAGTG   247

179  luLeuTyrThrSerAspThrValMetGlnAsnProGlnPheIleLysAla  195
     ||||||||||||||||||||||||||||||||||||||||||||||||||
248  AGCTGTACACCAGTGATACTGTCATGCAGAACCCACAGTTCATCAAAGCC   297

196  ThrIleValHisGlnAspGlnAlaTyrAspAspProLysIleTyrTyrPhePh  212
     ||||||||||||||||||||||||||||||||||||||||||||||||||
298  ACCATCGTGCACCAAGACACCAGCTTACGATGACAAGATCTACTACTTCTT  347

212  eArgGluAspGlnAsnProAspLysAsnProGluAlaProLeuAsnValSerA  229
     ||||||||||||||||||||||||||||||||||||||||||||||||||
348  CCGAGAGGACAATCCTGACAAGAATCCTGAGGCTCCTCTCAATGTGTCCC   397

229  rgValAlaGlnLeuCysArgGlyAspGlnGlyGlyGluSerSerLeuSer  245
     ||||||||||||||||||||||||||||||||||||||||||||||||||
398  GTGTGGCCCAGTTGTGCAGAGGGGACCAGGTGGGAAAGTTCACTGTCA    447

246  ValSerLysTrpAsnThrPheLeuLysAlaMetLeuValCysSerAspAl  262
     ||||||||||||||||||||||||||||||||||||||||||||||||||
448  GTCTCCAAGTGGAACACTTTTCTGAAAGCCATGCTGGTATCGCAGTGATGC  497

262  aAlaThrAsnLysAsnPheAsnArgLeuGlnAspValPheLeuLeuProA  279
     ||||||||||||||||||||||||||||||||||||||||||||||||||
498  TGCACCAACAAGAACTTCAACAGGCTGCAAGACGTCTTCCTGCTCCCTG    547

279  spProSerGlyGlnTrpArgAspThrArgValTyrArgValValPheSerAsn  295
     ||||||||||||||||||||||||||||||||||||||||||||||||||
548  ACCCCAGCGGCCAGTGGAGGGACACCAGGGTCTATGGTGTTTCTCCAAC    597

296  ProTrpAsnTyrSerAlaValCysValTyrSerLeuGlyAspIleAspPLy  312
     ||||||||||||||||||||||||||||||||||||||||||||||||||
598  CCCTGGAACTACTCAGCCGTGTGTGTATTCCCTGGGTGACATTGACAA    647

312  sValPheArgThrSerSerLeuLysGlyTyrHisSerSerLeuProAsnP  329
     ||||||||||||||||||||||||||||||||||||||||||||||||||
648  GGTCTTCCGTACCTCCTCACTCAAGGGCTACCACTCAAGCCTTCCCAACC   697

329  roArgProGlyLysCysLeuProAspGlnGlnProAspThrGluThr    345
     |||||||||||||||||||||||||||||||||||||||||:::
698  CGCGGCCTGGCAAGTGCCTCCCAGACCCAGCCAGCCGATACCCCACCACA   747
```

seq_name: N_Geneseq_36:X02664

seq_documentation_block:
ID   X02664   standard; DNA; 3999 BP.
AC   X02664;
DT   07-MAY-1999 (first entry)
DE   EP-892047 Seq ID 40.
KW   Semaphorin L; human; immunosuppressant; anti-inflammatory; gene therapy;
KW   organ transplantation; inflammation therapy; immunotherapy; agonist;
KW   immunomodulatory; antagonist; ss.
OS   Homo sapiens.
PN   EP-892047-A2.
PD   20-JAN-1999.
PF   06-JUL-1998; 112470.
PR   11-FEB-1998; DE-005371.
PR   09-JUL-1997; DE-029211.
PA   (HMRI) HOECHST MARION ROUSSEL DEUT GMBH.

PT   Ensser A, Fleckenstein B;
DR   WPI; 99-083564/08.
PT   New semaphorin L proteins - used as immunosuppressants and
PT   antiinflammatory agents in organ transplants, inflammation therapy,
PT   immunotherapy and gene therapy
PS   Disclosure; Page 103-105; 135pp; German.
CC   This invention describes a novel human semaphorin L protein. This protein
CC   or its encoding DNA are useful as immunosuppressants and/or
CC   anti-inflammatory agents in organ transplantation, inflammation therapy,
CC   immunotherapy and gene therapy. The DNA can be used to produce knock-out
CC   or knock-in animals for research purposes. The proteins or DNA can be
CC   used to search for the corresponding receptors or to screen for
CC   immunomodulatory agonists or antagonists.
SQ   Sequence 3999 BP;   1001 A;   986 G;   1021 C;   991 T;

alignment_scores:
     Quality: 1049.00   Length: 189
     Ratio: 5.550   Gaps: 0
Percent Similarity: 100.000   Percent Identity: 98.942

alignment_block:
US-09-041-236-2 x X02664

Align seg 1/1   to: X02664   from: 1   to: 3999

```
446  TyrValSerSerGlnTrpGluValSerGlnValProLeuAspLeuCysGl  462
     ||||||||::::||||||||||||||||||||||||||||||||||||||
154  CATGCGGAGCTCCCAGTGGGAGGTGAGCCAGGTGCCCCTGGACCTGTGTGA  203

462  uValTyrGlyGlyCysHisGlyCysSerHisGlyCysSerLeuMetSerArgArgAspProTyrC  479
     |||||||||||||||||||||||||||||||||||||||||||||||||
204  GGTCTATGGCGGGGGCTGCCACGGTGCCCTCATGTCCCGAGACCCCTACT   253

479  ySGlyTrpAspGlnGlyArgCysIleSerSerIleTyrSerSerGluArgSer  495
     |||||||||||||||||||||||||||||||||||||||||||||||||
254  GCGGGCTGGGACCAGGGCCGCTGCATCTCCATCTACAGCTCCGAACGGTCA  303

496  ValLeuGlnSerIleAsnProAlaGluProHisLysGlyCysProAsnPr  512
     |||||||||||||||||||||||||||||||||||||||||||||||||
304  GTGCTGCAAATCCATTAATCCAGCCGAGCCACACAAGGAGTGTCCCAACCC  353

512  oLysProAspLysProLeuGlnLysValSerLeuAlaProAsnSerA  529
     |||||||||||||||||||||||||||||||||||||||||||||||||
354  CAAACCAGACAAGCCCCACTGCAGAAGTTCCCTGGCCCCAAACTCTC   403

529  rgTyrTyrLeuSerCysProMetGluSerArgHisAlaThrTyrSerTrp  545
     |||||||||||||||||||||||||||||||||||||||||||||||||
404  GCTACTACCTGAGCTGCCCCATGGAATCCCGCCACGCCACCTACTCATGG   453

546  ArgHisLysGluAsnValGluGlnSerCysGluProGlyHisGlnSerPr  562
     |||||||||||||||||||||||||||||||||||||||||||||||||
454  CGCCACAAGGAGAACGTGGAGCAGAGCTGCGAACCTGGTCACCAGAGCCC   503

562  oAsnCysIleLeuPheIleGluAsnLeuThrAlaGlnGlnTyrGlyHisT  579
     |||||||||||||||||||||||||||||||||||||||||||||||||
504  CAACTGCATCCTGTTCATCGAGAACCTCACGGCGCAGCAGTACGGCCACT  553

579  yrPheCysGluAlaGlnGlyGlySerTyrPheArgGluAlaGlnHisTrp  595
     |||||||||||||||||||||||||||||||||||||||||||||||||
554  ACTTCTGCGAGGCCCAGGAGGGCTCCTACTTCCGCGAGGCTCAGCACTGG   603

596  GlnLeuLeuProGluProGluSerGlyIleLeuMetAlaGluHisLeuGlyHisAl  612
     |||||||||||||||||||||||||||||||||||||||||||||||||
604  CAGCTGCTGCCCGAGGACGGACATCATGGCCGAGCACCTGCTGGTCATGC  653

612  aCysAlaLeuAlaAlaSerLeuTrpLeuGlyValLeuProThrLeuThrL  629
     |||||||||||||||||||||||||||||||||||||||||||||||||
654  CTGTGCCCTGGCTGCCTCCCCTCCCCTGCTGGGGGTGCTGCCCACACTCACTC  703

629  euGlyLeuLeuValHis  634
     ||||:||||||||||
704  TTGCCATTGCTGTCCAC  720
```

```
seq_name: N_Geneseq_36:Q87442

seq_documentation_block:
ID   Q87442 standard; cDNA; 2601 BP.
AC   Q87442;
DT   21-NOV-1995  (first entry)
DE   Human semaphorin III cDNA.
KW   Semaphorin; grasshopper; human; vaccinia virus; Drosophila; Tribolium;
KW   variola major virus; smallpox; semaphorin receptor binding activity;
KW   modulation; nerve cell growth; immune response; viral pathogenesis;
KW   neurological disease; neuro-regeneration; oncological infection; ds.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   cds             16..2331
FT                   /tag= a
FT                   /product= human semaphorin III
PN   WO9507706-A.
PD   23-MAR-1995.
PF   13-SEP-1994; U10151.
PR   13-SEP-1993; US-121713.
PA   (REGC ) UNIV CALIFORNIA.
PI   Bentley DR,  Goodman CS,  Kolodkin AL,  Matthes D;
PI   O'Connor T;
DR   WPI; 95-131177/17.
DR   P-PSDB; R71380.
PT   New class of semaphorin peptide(s) and polypeptide(s) - are
PT   potent modulators of nerve cell growth and regeneration
PS   Example 2; Page 60-63; 101pp; English.
CC   The sequence of the cDNA encoding the human semaphorin III protein.
CC   The proteins encoded by the grasshopper semaphorin I (Q87441), human
CC   semaphorin III, vaccinia virus semaphorin IV (Q87443), Drosophila
CC   semaphorin I and II (Q87444-5), Tribolium semaphorin I (Q87446) or
CC   variola major (smallpox) virus semaphorin IV (Q87447) genes were used to
CC   generate a series of peptides (R70370-R70418), which retain semaphorin
CC   receptor binding activity. The semaphorin derived or semaphorin receptor
CC   derived peptides are potent modulators of nerve cell growth, immune
CC   responsiveness and viral pathogenesis.  They can be used in diagnosis and
CC   treatment of neurological disease and neuro-regeneration, immune
CC   modulation and diagnosis and treatment of viral and oncological infection
CC   and diseases.
SQ   Sequence  2601 BP;  809 A;  533 C;  593 G;  666 T;

alignment_scores:
     Quality:      649.50         Length:       687
     Ratio:          1.620        Gaps:          27
Percent Similarity:  58.370    Percent Identity:  28.093

alignment_block:
US-09-041-236-2  x  Q87442

Align seg 1/1  to: Q87442  from: 1  to: 2601

  1 LeuLeuLeuLeuTrpAlaAlaAlaAlaaSerAlaGlnGlyHisLeuAr 17
    :::|||| ::|||||:::|||||:::|||:|||||||
 34 ATTGTCTGTCTTTTCGGGGAGTATTACTTACAGCAAGAGCAAACTATCA 83

 17 gSerGly........ProArgIlePheAlaValTrpLysGlyHisV 30
    :::|||||:::         |||||||:::|||||    ::|||
 84 GAATGGGAAGAACAATGTGCCAAGGCTGAAATTATCCTACACAAGAAATGT 133

 30 alGlyGlnAspArgValAsp............PheGlyGlnThrGlu 41
    |:::||||||:::|||        :::::||||||:::||||
134 TGGAATCCAACAATGTGATCACTTTCAATGGGCTTGGCCAACAGCTCCAGT 183

 42 ProHisThrValLeuPheHisGluProGlyGlySerSerValTrpValGl 58
    :|||||:::|||    ||  ::||| :::::|||||||
184 TATCATACCTTCCTTTTGGATGAGGAA...CGGAGTAGGCTGTATGTTGG 230

 58 yGlyArgGlyLysValTyrLeuPheAspPheProGluGlyLysAsnAlaS 75
    |:::||:::|||||||  :::|||:::  :::|||::|||
231 AGCAAAGGAATCACATATTTTCATTCGACCTGGTTAATATCAAGGATTTTC 280
```

```
 75 erValArgThrValAsnIleGlySerThrLys.............. 85
    ::     ||:::|        ::::::|||:::
281 AAAAGATGTGTGCCAGTATCTTACACCAGAGAAGAATGAATCAAGTGG 330

 86 ...GlySerCysLeuLysArgAspCysGluAsnTyrIleThrLeuLe 101
       |||  :::||| ::::||| :::|||:::||| :::|||
331 GCTGGAAAGACATCCTGAAA...GAATGCTAATTTCATCAAGGTACT 377

101 uGluArgArgSerGlu...GlyLeuLeuAlaCysGlyThrAsnAlaArgH 117
    |||||  ||||||      :::||||||::|||||||:::||||
378 TAAGGCATATAATCAGACTCACTTGTACGCCTGTGGAACGGGGCTTTC 427

117 lsProSerCysTrpAsnLeuValAsnGlyThrValValProLeuGlyGlu 133
    ||||| |||       :::      :::||| ::::::||||||
428 ATCCAATTGC.............ACCTACATTGAAAATTGGACAT 459

134 Met........................................... 134
    |||
460 CATCCTGAGGACACAATATTTTAAGCTGGAGAACTCACATTTGAAAACGG 509

135 .ArgGlyTyrAlaProPheSerProAspGluAsnSerLeuValLeuPheG 151
     ||||:::  |||||:::||| :::||| :::|||:::|||   :::||
510 CCGTGGGAAGAGTCCATATGACCCTAAGCTGCTGACAGCATCCCTTTAA 559

151 luGlyAspGluValTyrSerThrIleArgLysGlnLeuTyrAsnGlyLys 167
    ||||||||||:::|:::|||||               :::
560 TAGATGGAGAATTATACTCTGGAACTGCAGCTGATTTTATGGGGCCGAGAC 609

168 IleProArgPheArgIleArgGlyGluSerGluLeuTyrThr....... 182
    |||||||          ||||||         :::    |||
610 TTTGCTATCTTCCGAACTCTGGGCACCACCCAATCAGGACAGAGCA 659

183 ....SerAspThrValMetGlnAsnProGlnPheIleLysAlaThrIlev 198
        :::      |||:::|||||||:::||||||:::|||||||  ::
660 GCATGATTCCAGGTGGCTCAATGATCCAAAGTTCATTAGTGCCCACCTCA 709

198 alHisGln...AspGlnAlaTyrAspAspLysIleTyrTyrProPheArg 213
    ::  :::   |||||:::|||||:::||| |||||||||||||||||
710 TCTCAGAGAGTGACAATCCTGAAGATGACAAAGTATACTTTTTCTTCCGT 759

214 GluAspAspProLysAsnProGluAlaProLeuAsnValSerArgVa 230
    ||||:::      :::   ::::::|||::|||||||::::::|||
760 GAAAATGACAATAGGAACACTCTGAAAAGCTACTCACCGTAGAAT 809

230 lAlaGlnLeuCysArgGlyAspGlnInTrpGlyGlyGluSerLeuSerVals 247
    |  ::::::|:::||||:::|||||||:::|||:::|||||||
810 AGGTCAGATATGCAAGAATGACTTTGGAGGGCCACAGAAGTCTG...GTGA 856

247 erLysTrpAsnThrPheLeuLysAlaMetGlnLysCysSerSerAspAlaAla 263
    :::||||||:::|||||:::|||||||     |||||:::|||
857 ATAAATGGACAACATTCCTCAAAGCTCGTCTGATTTGCTCAGTGCCAGT 906

264 ThrAsn..........LysAsnPheAsnArgLeuGlnAspValPheLeuLe 277
    :::|||          ||:::||| :::|||  :::||||||
907 CCAAATGCCATTGACACACCATTTTGATGAACTGCAGGATGTAATTCCTAAT 956

277 u......ProAspProSerGlyGlnTrpTrpArgAspThrArgGlyLyV 292
           ||||||:::||| :::   |||:::||| |||||||
957 GAACTTTAAAGATCCT..........AAAAATCCAGTTGTATATGGAG 994

292 alPhe..........SerAsnProTrpAsnTyrSerAlaValCysValTyr 305
    |||||          :::|||||||  :::      |||||:::
995 TGTTTACGACTTCCAGTAACATTTTCAAGGGATCAGCCGTGTGATGTAT 1044

306 SerLeuGlyAspIleAspLysValPhe............... 314
    :::|||||||:::|||   |||||||
1045 AGCATGAGTGATGTGAGAAGGGTTCCTTGGTCCATATGCCCCACAGGGA 1094

315 ....ArgThrSerSerLeuLysGlyTyrHisSerSerLeuProAsnProA 330
        :::||||||:::|||  :::|||||||:::|||:::||| ::||:::
1095 TGGACCCAACTATCAATGGGTGCCTTATCAAGGAAGAGTCCCCTATCCAC 1144
```

```
                                                       ::|||||||||
                                              2024 CAGAGCATTTG 2034

                                         seq_name: N_Geneseq_36:T60666

                                         seq_documentation_block:
                                         ID   T60666 standard; cDNA; 4391 BP.
                                         AC   T60666;
                                         DT   24-JUL-1997  (first entry)
                                         DE   Mouse CD100 antigen cDNA.
                                         KW   CD100 antigen; semaphorin; leukocyte; B cell; T cell; lymphocyte;
                                         KW   vaccine; ss.
                                         OS   Mus sp.
                                         FH   Key              Location/Qualifiers
                                         FT   cds              434..3019
                                         FT                    /*tag= a
                                         FT   signal_peptide   434..556
                                         FT                    /*tag= b
                                         FT   mat_peptide      557..3016
                                         FT                    /*tag= c
                                         PN   WO9717368-A1
                                         PD   15-MAY-1997.
                                         PF   12-NOV-1996; U18645.
                                         PR   09-NOV-1995; US-556422.
                                         PA   (DAND ) DANA FARBER CANCER INST.
                                         PI   Boussiotis V,  Freeman GJ,  Hall KT,  Nadler LM,  Schultze JL;
                                         DR   WPI; 97-280982/25.
                                         DR   P-PSDB; W17658.
                                         PT   Nucleic acid molecule encoding CD100 antigen - which stimulates
                                         PT   leukocyte response, e.g. B cell aggregation, differentiation,
                                         PT   survival and T cell proliferation
                                         PS   Example 8; Page 81-85; 135pp; English.
                                         CC   A cDNA molecule (T60666) codes for mouse CD100 antigen (W17658) a
                                         CC   novel leukocyte semaphorin-like protein that stimulates a leukocyte
                                         CC   response, including B cell aggregation, B cell differentiation, B
                                         CC   cell survival and/or T cell proliferation.  It was isolated from a
                                         CC   a cDNA library prepd. from murine t cells activated with anti-CD3
                                         CC   and anti-CD28 antibodies using a probe corresponding to the
                                         CC   semaphorin and Ig domains of the human CD100 sequence (see also
                                         CC   T60666).  CD100 nucleic acids, antisense sequences, recombinant
                                         CC   vectors, host cells expressing CD100, and CD100 polypeptides can
                                         CC   be utilised in diagnostic and therapeutic methods involving
                                         CC   modulation cf B and T cell responses, neuron axonal growth and
                                         CC   immune cell-nerve cell interaction.
                                         SQ   Sequence 4391 BP;  1009 A;  1180 C;  1105 G;  1097 T;

                                         alignment_scores:
                                              Quality:    594.00          Length:    695
                                              Ratio:      1.543           Gaps:       32
                                         Percent Similarity: 55.396    Percent Identity: 28.201

                                         alignment_block:
                                         US-09-041-236-2 x T60666

                                         Align seq 1/1  to: T60666  from: 1  to: 4391

                                            1 LeuLeuLeuLeuTrpAlaProAlaAlaAlaAlaSerAlaGlnGlyHisLeuAr 17
                                              |||::|||||||::|   :|||||||||||||||||  |||
                                          473 CTGGTGGTAGTGTTGAAGAACCGCGGTGGCATTTGCA............... 508

                                           17 gSerGlyProArgIlePheAlaValTrpLys...GlyHisValGlyGlnA 33
                                              |||||:::|||||:::::                ||:|||||||
                                          509 .CCTGTGCCTCGGCTC..........ACCTGGGAACATGGAGAGGTAGGT... 547

                                           33 spArgValAspPheGlyGlnThrGluProHisThrValLeuPheHisGlu 49
                                              :||::  |||  ::|||||  |||||||||||:| :::|||||
                                          548 .................................CTGGTGCAGTTCACAAG 565

                                           50 ProGly................................SerSerSe 54
                                              ||||||                                :
                                          566 CCAGGCATCTTAACTACTCGGCCTTGCTGATGAGTGAGGACAAAGACAC 615

                                          605 laGluHisLeu 608
```

```
330 rgProGlyLysCys..........................LeuProAspGlnGln 339
    ||||||||                               :::  ||  |||
1145 GGCCAGGAACTTGTCCCAGCAAAACATTGGTGGTTTGACTCTACAAAG 1194

340 ProIleProThrGluThrPheGlnValAlaAspArgHisProGluValAl 356
    :::||||   ||:|||||::               |||||   :::
1195 GACCCTCCTGATGATGTTATAACCTTTGCAAGAAGTCATCCAGCCATGTA 1244

356 aGlnArgValGluProMetGly...........ProLeuLysThrProL 369
    |||||||||||||||||                ||||||| ||
1245 CAATCCAGTGTTCCTATGAACAATCGCCCAATAGTGATCAAAACGGGATG 1294

369 euPheHisSerLysTyrHisTyrGlnLysValAlaValHisArgMetGln 385
    ::  ||:|||||||||||   ||||::|| |||:::::||
1295 TA.......AATTATCAATTACACAAATTGTCGTAGACCGAGTGGAT 1335

386 AlaSerHisGlyGluThrPheHisValLeuTyrLeuThrThrAspArgGl 402
    |||:::||||   ||||||::  :::|||||| |||||  |
1336 GCAGAGATGACAG...TATGAGTTAATGTTATCGGACAGATGTTGG 1382

402 yThrIleHisLysValValGlu...ProGlyGluGlnHisSerSerPheA 418
    |||||:::|||||  |||:::   ||||||||||   :::|||
1383 GACCGTTCTTAAGTAGTTTCAATTCCTAAGGAGACTTGGTATGATTTAG 1432

418 laPheAsnIleMet...GluIleGlnProPheArgArgAlaAlaAlaIle 433
    :::||||||       ||||||   ||||||||||   :::|||||
1433 AAGAGGTTCTGCTGGAAGAAATGACAGTTTTCGGGAACCGACTGCTATT 1482

434 GlnThrMetSerLeuAspAspAlaGluArgArgLysLeuTyrValSerSerGl 450
    ::::||||||:::||||||||:::||||||||||||||||||||||:::
1483 TCAGCAATGGACGTTTCCACTAAGCAGCCAACTAATATATTGGTTCAAC 1532

450 nTrpGluValSerGlnValProLeuAspLeuCysGluValTyrGlyGlyG 467
    :::||||||::|||||   ::::||||||:::|||||||   ||||||
1533 GGCTGGGGTTGCCCAGTCCCCTTTACACCCGTGTGATATTTACGGGAAAG 1582

467 lyCysHisGlyGlyLeuMetSerArgAspProTyrCysGlyTrpTrpAspGln 483
    ::||   |||||||||||:::||||||||||||||||||:::||||||
1583 CGTGTGCTGAGTGTGTGCCTCGCCCGAGACCCTTACTGTGCTTGGGATGGT 1632

484 GlyArgCysIleSerIleTyr.....SerSerGluArgSerValLeuGl 498
    :::|||   |||   ::                ::|||:::||||
1633 TCTCGATGTTCTCGCTATTTCCCACTGCCAAGAGACGCCACAAGACGACA 1682

498 nSerIleAsnProAlaGluProHisLysGluLysCysProAsnProLysProA 515
    ||::::|||   :::||||||||||:::::||||||||::|||
1683 AGATATAAGAAATGGAGACCCACTGACTCACTGTTCAGACTTACACCATG 1732

515 spLys..........AlaProLeuGlnLysValSerLeu...AlaPro 526
    :||:           :::||||                   :::
1733 ATAATCACCATGGCCACAGCCTGAAGAGAGAATCATCATGGTGTAGAG 1782

527 AsnSerArgTyrTyrLeuSerCysProMetGluSerArgHisAlaThrTy 543
    ||||||   :::||:||||||||||||||||||||||||
1783 AATAGTAGACACATTTTGGAATGCAGTCCGAAGTCCCAGAAGCGCCTGGT 1832

543 rSerTrpArgHisLysGluAsnValGluGlnSerCysGluProGlyHisG 560
    :::||||||||:::::||||||||||::|||||||||::
1833 CTATTGGCAATTCCAGAGGCGAAATGAACGAGAAGAGATCAGAG 1882

560 lnSerProAsnCysIle.......LeuPheIleGluAsnLeu 571
    ::||||||:::|||            |||||||::::||||
1883 TGGATGATCATATCATCAGGACAGATCAGGCCTTCTGCTACGTAGTCTA 1932

572 ThrAlaGlnGlnTyrGlyHisTyrPheCysGluAlaGlnGluGlySerTy 588
    :::||||||||:::||:::|| |||  :::|||||||
1933 CAACAGAAGGATTCAGGCAATTACCTCTGCCATGCGGTGGAACATGGGTT 1982

588 rPheArgGluAlaGlnHisTrpGlnLeuLeuProGluAspGlyIleMetA 605
    :::||||||   :::||                        :::||
1983 CATACAAACTCTTCTTAAGGTAACCCTG........GAGTCATTGACA 2023

605 laGluHisLeu 608
```

```
                         |||                    :::..
1468 TGCCACAGTGGACAGTCTCACCACCAGTGGGTCGCGTTACAATGGCCCAG 1517

326 euProAsnProArgProGlyLysCysLeuPheAspGlnGln.........  339
    |||:||||||||||||                      |||  |||:
1518 TGCCCACTCCCCGACCTGGAGCGTGTATCGACAGTGAGGCCCGGCAGCC   1567

340 .............ProIleProThrGluThrPheGlnValAlaAlas     350
                  ::|||       |||:|| ||:|||
1568 AACTACACCAGTCACCTTGAATCTCCCAGACAAAACACTGCAGTTGTAAA 1617

350 pArgHisProGluValAlaGlnArgValGlyProMetGlyProLeuLysT 367
    ||||||                  |||:        :::
1618 AGACCACCCTTTGATGGAGCCTCAGTGACCCGATAGACAAC...AGAC   1664

367 hrProLeuPheHisSerLysTyrGlnTyrHisTyrGlnLysValAlaValHisValArg 383
    |||::|                      :::
1665 CCAAGCTGATCAAAAAAGATGTAAACTACACCCAGATAGTGGTAGACAGG 1714

384 MetGlnAlaSerHisGlyGluThrPheHisIleValLeuTyrLeuThrThrAs 400
    :||||||                 |:|        ||||||:::|||||
1715 ACCCAGGCCCGGATGGGGGACTTTCTACGACGTCATGTTCATCAGCACAGA 1764

400 pArgGlyThrIleHisLysValValGluProGlyGluGlnHisSerSerP 417
    |||||::||:::||||::|||                :::      |||
1765 CCGGGGAGCTCTGCTATCGTCAAAGACAGCATCATCCTCAAAAGAGGTCATGTC. 1813

417 heAlaPheAsnIleMetGluIleGlnProPheArgArgAlaAlaAlaAlaIle 433
    ..........   ||||                      :::
1814 ..ATCGAGGAGACCCAACTCTTCCGGACTCTGAACCGGTC 1852

433 GlnThrMetSerLeuAspAlaGlu......ArgArgLysLeuTyrValSe 448
    |||:|||:::::::|||         |||
1853 CTAACTCGTGTCTATCGTCAAAGAGGGAGGAAGTTTGTCTATGCAGG   1902

448 rSerGlnTrpGluValSerGlnValProLeuAspAspLeuCysGluValTyrG 465
    ::|||:||          ||||||||:|||||||
1903 CTCCAACTCTGGAGTGGTCCAAGCCGCCCTGGCATTCTGCGAAAAGCAC. 1951

465 lyGlyGlyCysHisGlyCysLeuMetSerSerArgAspProTyrCysGlyTrp 481
    ||||:||||:||:||||                :::
1952 .GGTAGCTGTGAAGACTGTGTTGTTAGCACGGACCCCTACTGTGCCTGG 1999

482 AspGlnGly......ArgCysIleSerIleTyr.......SerSerGl   493
    ||::||         ::|||                ||||
2000 AGCCCAGCCATCAAGGCCTGTGTTACCCTGCACCAGGAGAGGCCTCCAG 2049

493 uArgSerValLeuGlnSerIleAsnProAlaGluProHisLysGluCysP 510
    :|||||::|||          |||:::||||||||
2050 CAGGGGCTGGATTCAGGACATGAGC............            2074

510 roAsnProLysProAspLysProLeuGluLysValSerLeuAlaPro   526
    |||:::||          ||:::           :::
2075 .GGTGACACATCCTCATGCCTGGATAAGAGTAAAGAAAGTTTC      2116

527 AsnSerArgTyrTyr...................LeuSerCysProMe   536
    |||:::                 :::
2117 AACCAGCATTTTTCAACGCACCGGCCAACTCAAATGTTTCCA       2166

536 tGluSerArgHisAlaThrTyrSerTrpArgHisLysGluAsnValGluG 553
    ::||:||        :::|||          ||||||
2167 AAAGTCCAACCTAGCCCGGTGGTATGGAAGTTCCAGAATGGCGAGTTGA  2216

553 lnSerCysGluPro..................GlyHisGlnSerProAsnCys 564
    |||:::|||             ||:::
2217 AGGCCCGAAGTCCCAAGTACGGCTTTGTGGGCAGGAAGCAC        2257

565 IleLeuPheIleGluAsnLeuThrAlaGlnIleThrArgGlyHisTyrPheCy 581
    ||||:|:||                ||| ||:::
2258 ..CTGCTCATCTTCAACCTGTCGGACGCAGCGGCGTGTACCAGTG      2304

581 sGluAlaGlnGlyGlyLysTyrPheArgGlyPheArgGluGlnHisTrpGlnHisLeuL 598
    |   |||||||               |||||||
```

```
2305 CCTGTCAGGAA......AGGGTGAGGAATAAAACGGTCTCCCAGCTGC 2348

 598 euProGluAspGlyIleMetAlaGluHisLeuLeu 609
      ||    |||::::|||:::|||
2349 TG..............GCCAAGCACGTTCTG 2365

seq_name: N_Geneseq_36:X03792

seq_documentation_block:
ID   X03792 standard; cDNA; 2975 BP.
AC   X03792;
DT   31-MAR-1999 (first entry)
DE   Human semaphorin E encoding cDNA clone BR5334.
KW   Human; semaphorin E; clone BR5334; nutritional; immune stimulating;
KW   vaccine; haematopoiesis regulating activity; tissue growth;
KW   haemostatic; thrombolytic; anti-inflammatory; tumour inhibition;
KW   gene therapy; ds.
OS   Homo sapiens.
FH   Key            Location/Qualifiers
FT   CDS            245..2500
FT                  /*tag= a
PN   WO9853065-A1.
PD   26-NOV-1998.
PF   19-MAY-1998; U10188.
PR   18-MAY-1998; US-080695.
PR   19-MAY-1997; US-858834.
PA   (GEMY ) GENETICS INST INC.
PI   Agostino MJ, Fechtel K, Howes SH, Jacobs K, Lavallie ER,
PI   McCoy JM, Merberg D, Racie LA, Spaulding V, Treacy M;
DR   WPI; 99-059742/05.
DR   P-PSDB; W30617.
PT   New polypeptides encoding secreted human proteins - derived from a
PT   human foetal kidney cDNA library
PS   Claim 1: Page 43-44: 58pp; English.
CC   The present sequence encodes human semaphorin E from cDNA clone BR5334.
CC   Human semaphorin E polynucleotide sequences and protein sequences from
CC   the present invention, are predicted to have biological activities which
CC   would make them suitable for treating, preventing or ameliorating
CC   medical conditions in humans and animals, although no supporting data is
CC   given. Suggested activities include nutritional activity, immune
CC   stimulating (e.g. as vaccines) or suppressing activity, haematopoiesis
CC   regulating activity, tissue growth activity, activin/inhibin activity,
CC   chemotactic/chemokinetic activity, haemostatic and thrombolytic
CC   activity, receptor/ligand activity, anti-inflammatory activity,
CC   cadherin/tumour invasion suppressor activity, and tumour inhibition
CC   activity. The polynucleotide sequences are also stated to be useful for
CC   gene therapy. A host cell transfected with the polynucleotide sequence
CC   encoding human semaphorin E or its subfragments and variants is useful
CC   for recombinant production of the clone BR5334 related protein.
SQ   Sequence 2975 BP;   887 A;   620 C;   665 G;   802 T;

alignment_scores:
     Quality:      591.50       Length:    610
     Ratio:        1.666        Gaps:       24
Percent Similarity:  58.197    Percent Identity:  29.016

alignment_block:
US-09-041-236-2 x X03792

Align seg 1/1  to: X03792  from: 1  to: 2975

  43 HisThrValLeuPheHisGluProGlySerSerValTrpValGlyGl 59
      ::  ||:::|||::::||     ::::|||     :::|||:::
 404 TACAGGATTTTATTAATGGATGAAGATCAGGACCGGATATATGTGGGAAG 453

  59 yArgGlyValTyrLeuPheAspPheProGluGlyLysAsnAlaSer. 75
      ::|||         ::|||         :::|||         |||
 454 CAAAGATCACATTCTTTCCCTGAATATT........AACAATATAAGTC 494

  76 ..........VarArgThrValAsnIleGlySer 83
                ||| ::::::|||:::::||: :::
 495 AAGAAGCTTGAGTGTTTCTGGCCAGCATCCAATCAACATCAAAGTTGAAGAA 544
```

```
  84 ThrLysGlySerCysLeuAsp..LysArgAspCysGluAsnTyrIleTh 99
      |||  :::   |||  ||| |||  ||:::|||
 545 TGCAAAATGGCTGGCAAAGATCCCACACACGGCTGTGGGAACTTTGTCCG 594

  99 rLeuLeuGluArg...ArgSerGluGlyLeuLeuAlaCysGlyThrAsnA 115
      :::||     ||||||   ||| |||:::|||   :::||||||:::|||
 595 TGTAATTCAGAGCTTTCAATCCACACATTGTATGTCGTGGGAGTGGCG 644

 115 laArgHisProSerCysTrpAsnLeuValAsnGly............ 126
      ||     |||:::|||   |||   ::|||
 645 CTTTCAGTCCTGTCTGTACTTACTTGAACAGAGGAGGAGATCAGAGGAC 694

 127 ThrValValProLeuGlyGluMet...........ArgGlyTyrAlaPr 139
      |||     :::|||   :::         ::|||     :::
 695 CAAGTTTCATGATTGACTCCAAGTGTGAATCTGGAAAAGGACCGCTGCTC 744

 139 oPheSerProAspGluAsnSerLeuValLeuPheGluGlyAspGluValT 156
      |||:::||:::|||     ::::|||||     :::::||||||:::|||
 745 TTTCAACCCCAACGTGAACACGGTGTCTGTTATGATCAATGAGGAGCTTT 794

 156 yrSerThrIleArgLysGlnGluTyrAsnGlyLysIleProArgPheArg 172
      ::::||         :::        :::|||             |||||
 795 TCTCTGGAATGTATATAGATTTCATGGGGACAGATGCTGTCTATTTTTCGA 844

 173 ArgIleArgGlyGluSerGluGluTyrThr.........SerAspThrVa 186
      :::         :::  ||: ::: |||             |||:::|||
 845 AGTTTAACCAAGAGGAATGCGGTCAGAACTGATCAACATAATTCCAAATG 894

 186 lMetGlnAsnProGlnPheIleLysAlaThrIleValHisGln...AspG 202
      :::|||   :::|||           :::|||   :::         ||:
 895 GCTAAGTGAACCTATCTTTGTAGAATGCACATGTCATCCCAGATGGTACTG 944

 202 lnAlaTyrAspAspLysIleTyrPheArgGluAspAsnProAsp 218
      ::         :::|||:::|||::||:::|||::::|||||||
 945 ATCCAAATGATGCTAAGGTGACTTCTTCTTCAAAGAAAAACTGACTGAC 994

 219 LysAsnProGluAlaProLeuAsnValSerArgValAlaGlnLeuCysAr 235
      :::|||           |||   ||::||:::|||   ::||::::|||
 995 AATAACAGGAGCACGAAACAGATTCATTCCATGATTGCTCGAATATGTCC 1044

 235 gGlyAspGlnGlyGlySerLeuSerValSerLysTrpAsnThrP 252
      :::|||::|||:::||||||:::|||:::|||||||
1045 TAATGACACTGGTGGACTGCCTAGCCTT...GTCAACAAGTGGACCACTT 1091

 252 heLeuLysAlaMetLeuValCysSer........AspAlaAlaThrAsn 265
      |||||||||     ||||||||||         ::::|||
1092 TCTAAAGGCGAGGCTGGTGTGCTCCGGTAACAGATGAAGACGGCCCAGAA 1141

 266 LysAsnPheAsnArgLeuGlnAspValPheLeuLeuProAspProSerGl 282
      :::|||:::|||||||||||||||||||||||::
1142 ACACACTTTGATGAATTAGAGGATGTGTTTCTGCTGGAAACTGATAACCC 1191

 282 yGlnTrpArgAspThrArgValTyrGlyValPhe.........SerAsnP 296
      ::||||||||     :::|||       ||::::::
1192 G......AGGACAACACTAGTGTATGGCATTTTACAACATCAAGCTCAG 1235

 296 roTrpAsnTyrSerAlaValCysValTyrSerLeuGlyAspIleAspLys 312
      ::::         ::::|||||||||||||||||||||||
1236 TTTTCAAAGGATCAGCCGTGTGTGTGTATCATTATCTGATATACAGACT 1285

 313 ValPheArg..........ThrSerSerLeuLy 320
      ||||||:::                ::: |||
1286 GTGTTTAATGGGCCTTTGCCCACAAAGAAGGGCCCAATCATCAGCTGAT 1335

 320 sGlyTyrHisSerSerLeuProAsnProArgProGlyLysCys......... 334
      :::|||||||     :::||||||||||
1336 TTCCTATCAGGGCAGAATTCCATCCTCCGCCTGAACTTGTCCAGGAG 1385

 335 ...........LeuProAspGlnProIleProThrGluThr 345
                  |||   |||   |||
1386 GAGCATTTCACCCAATATGCGAACCACCAAGAGTTCCAGATGATGTT 1435
```

```
OS   Homo sapiens.
FH   Key            Location/Qualifiers
FT   CDS            61..2646
FT                  /*tag= a
FT                  /product= "semaphorin"
PN   JI015490-A.
PD   16-JUN-1998.
PR   27-NOV-1996; 332900.
PF   27-NOV-1996; JP-332900.
PA   (SUMD ) SUMITOMO SEIYAKU KK.
DR   WPI; 98-391044/34.
DR   P-PSDB; W58540.
PT   New human semafolin gene - useful in the diagnosis of nervous system
PT   and immune disorders
PS   Claim 2; Page 12-13; 15pp; Japanese.
CC   The present sequence encodes human semaphorin (translated from the
CC   Japanese specification as semafolin). Semaphorin has nerve extension
CC   inhibiting activity. The semaphorin gene is useful for the diagnosis,
CC   treatment and researches on nervous diseases and immune diseases.
SQ   Sequence  2769 BP;    655 A;    780 C;    738 G;    596 T;

alignment_scores:
     Quality:     591.00       Length:       695
     Ratio:        1.539       Gaps:          32
     Percent Similarity: 55.252   Percent Identity: 28.201

alignment_block:
US-09-041-236-2 x V31121

Align seg 1/1  to: V31121  from: 1  to: 2769

    1 LeuLeuLeuLeuTrpAlaAlaAlaAlaSerAlaAlaGlnGlyHisLeuAr 17
      ||:||||||||:|||     :::|||:||||     ||
  100 CTGGTGGTAGTGTTGAGACCGCGGTGCCGCCATTGCA............... 135

   17 gSerGlyProArgIlePheAlaValTrpLys...GlyHisValGlyGlnA 33
      ||:|||||                       ||::|||     |||:|||
  136 .CCTGTGCCTCGGCTC..............ACCTGGGAACATGGAGAGGTAGGT..... 174

   33 spArgValAspPheGlyGlnThrGluProHisThrValLeuPheHisGlu 49
                                           ||  |||||||||::
  175 ..............................CTGGTGCAGTTTCACAAG 192

   50 ProGly...............................SerSerSe 54
      |||||||
  193 CCAGGCATCTTAACTACTCGGCCTTCGTGATGAGTGCAAAGACAC 242

   54 rValTrpValGlyArgGlyValGlyValTyrLeuPheAspPheProGluG 71
      ||:::||||||:|||||     ||||::||     ||||
  243 TCTGTATGTAGGCGCGCCCGGGAAGCAGTCTT................. 273

   71 lyLysAsnAlaSerValArgThrValAsnIleGly...............  82
                                     :::|||:||||||
  274 .............GCAGTGAATGCGGCTGAACATCTCGAGAAGCAACATGAG 312

   83 ...............................SerThrLysGlySerCysLeuAspLys..  91
                                     |||:||||||     :::|||
  313 GTATATTGGAAGGTCTCTGAGACAAAAAATCCAAGTGTGCAGAGAAGGG 362

   92 ...............ArgAspCysGluAsnTyrIleThrLeuLeuGlu...A 103
                    |||||||     ||::|||     :::|||
  363 GAAATCAAAGACGACGGAATGCCTAAACTACATTCGAGTACTACAGCCCAC 412

  103 rgArgSerGluGlyLeuLeuAlaCysGlyThrAsnAlaArgHisProSer 119
      |||||     :::||||||||     |||||||||||||||
  413 TAAGCACACTCCCCTATGTGTGGACCAATGCCGGTTCCAGCCCCACC...... 462

  120 CysTrpAsnLeuValAsnGlyThrValProLeuGlyGluMet........ 134
      |||:||||||     :::|||          ||||||
  463 TGTGACCACCTGAACTTGACATCTTCAAGTTTCTGGGGAAAAGTGAAGA 512

  135 ...ArgGlyTyrAlaProSerProAspGluAsnSerLeuValLeuP 150
```

```
  346 PheGlnAlaAlaAspArgHisProGluAlaAlaGlnArgValGluProMe 362
      ::|||||     |||     ::     |||:::
 1436 GTCACTTTTATTCGGAACCATCCTCCATGTACAATTCCATCTACCCAAT 1485

  362 tGlyProLeuLysThrHisT .......PheHisSerLysTyrHisT 376
      |||  |||||     ||:::          :::          |||
 1486 CCAC......AAAAGGCCTTGATTGTTCGTATTGGCACTACAAGT 1529

  376 yrGlnLysValAlaValHisArgMetGlnAlaSerHisGlyGluThrPhe 392
      ||  |||||||     ::||||||     |||||:::          ::
 1530 ACACAAAGATAGCTGTGGATCAGTGGACGCGCTGATGAGGGA...TAC 1576

  393 HisValLeuTyrLeuThrThrAspArgGlyThrIleHisLysValValGl 409
      |||||||||||||     ::::          ::|||:::|||||||||
 1577 CATGTCCTGTTTCTCGGAACAGTCGGGTACTGTGCAAAAAGTGGTTGT 1626

  409 u...ProGlyGluGlnGluHisSerPheAlaPheAsnIleMetGluIleG 425
      |||          |||  :::          ::     |||:::|||
 1627 TCTTCCTACTAACAACTCGTCAGTGGCGAGCTCATTCTGGAGGAGCTGG 1676

  425 lnProPheArgArgAlaIleAlaAlaIleGlnThrMetSerLeuAspAlaGlu 441
      |||          |||     :::|||:::          ::|||
 1677 AAGTCTTTAAGAATCATGCCTCCTATACAACAGTGAAAATTTCATCCAAA 1726

  442 ArgArgLeuTyrValSerSerGlnTrpGluValSerGlnValProLe 458
      :::|||||||:::     ||||||          |||:::|||  ||
 1727 AAGCAACAGTTGTATGTGAGTTCCAATGAAGGGGTTTCCCAGGTATCTCT 1776

  458 uAspLeuCysGluValTyrGlyGlyCysHisGlyCysLeuMetSerA 475
      |||:::||||||     |||  ||          ::|||:::
 1777 GCACCGCTGCCACATCTATGGTACAGCCTGTGCTGCTGCCTGCGC 1826

  475 rgAspProTyrCysGlyGlyTrpAspGlnGlyArgArgCysIleSerIleTyr... 490
      |||||||||||:||||||          ::|||
 1827 GGGACCCTTATTGCGGGTGATGGACGGATTCCAGAATCTCACCCA 1876

  491 ...SerSerGluArgSerValLeuGlnSerIleAsnProAlaGluProHi 506
      |||||          |||||||||     :::|||:::
 1877 ACTGGGAAACGGAGAGGCCGAGACAAGAGTGAGACATGAGAAACCCACT 1926

  506 sLysGlyCysPro........AsnProLysProAspLysProAlaProLeuGlnL 521
      ::|||          |||     |||:::
 1927 GACTCAATGCAGAGATTTAATCTAAAAGCATACAGAAATGCAGCTGAAA 1976

  521 ysValSerLeu...AlaProAsnSerArgTyrTyrLeuSerCysProMet 536
      ||||||          ::|||:::|||||||:::
 1977 TTGTGCAGTATGGAGTAAAAAAATAACAACCACTTTTCTGGAGTGTGCCCCC 2026

  537 GluSerArgHisAlaThrTyrSerTrp.................Ar 546
      |||:::          |||||||||                   ||
 2027 AAGTCTCCGCAGGCCATCTATCAAGTGGCTGTTACAGAAAGACAAAGACAG 2076

  546 gHisLysGlu...AsnValGluGlnSerCysGluProGlyHisGlnSerP 562
      |:::|||||               |||:::
 2077 GAGGAAAGAGGTTAAGCTGAATGAACGAATAATAGCCCACTTCACAGGGA 2125

  562 roAsnCysIleLeuPheGluIleGluAsnLeuThrAlaGlnInTyrGlyHis 578
      |||          |||:::          |||
 2126 ..........CTCCTGATCCGCCTCTTCAGGGTTCGACCAAGGACTT 2164

  579 TyrPheGluGlnValAlaGlyLySerTyr 588
      |||  |||||     :::|||
 2165 TATCACTGCAATTGCTACAGAAAATAGTTTC 2194
```

us-09-041-235-2.ЕЕ3.rng

Wed Jan 19 11:15:40 2000

```
1392 CCGGGGAGCTCGCATAAAGACAGTCATCCTTACAAAAGAGGTGCATGTC.  1440
 417 heAlaPheAsnIleMetGluIleGlnProPheArgArgAlaAlaAlaIle   433
                  ::::                  ||||||       ::  ::
1441 ..........ATCGAGGAGACCAACTCTTCCGGGACTTGAACCGGTC  1479

 434 GlnThrMetSerLeuAspAlaGlu.....ArgArgLysLeuTyrValSe   448
     ||| ::       ::: :::          :  :::  :  :: : :::
1480 CTAACTCTGCTGCTATCGTCAAAGAAGGGGAGGAAGTTTGTCTATGCAGG  1529

 448 rSerGlnTrpGluValSerGlnValProLeuAspLeuCysGluValTyrG   465
     :||:  |||       ||:|||||||  ::||||||       :::||||
1530 CTCCAACTCTGGAGTGGTCCAAGCGCCCCTGCATTCTGCGAAAAGCAC.  1578

 465 lyGlyGlyCysHisGlyCysSerLeuMetSerArgAspProTyrCysGlyTrp  481
     :||       :::::    ||||||||||:::::||||||:::
1579 ..GGTAGCTGTGAAGACTGTGTGTTAGCACGGACCCCTACTGTGCCTGG  1626

 482 AspGlnGly......ArgCysIleSerIleTyr........SerSerGl   493
     ||| :: |||||||  ||| ::  |||  :::        ::|||:
1627 AGCCCAGCCATCAAGGCCTGTGTTACCTGCACCAGGAAGAGGCCTCCAG  1676

 493 uArgSerValLeuGlnSerIleAsnProAlaGluProHisLysGluCysP   510
     :||:: ::: ||:|||:  :::||:||
1677 CAGGGGCTGGATTCAGGACATGAGC.  1701

 510 roAsnProLysProAspLysValProLeuGlnLysValSerLeuAlaPro   526
                ::: :::||:||||:::::||:||:::::||:::
1702 ........GGTGACACATCCTCATGCCTGGATAAGAGTAAAGAAAAGTTTC  1743

 527 AsnSerArgTyrTyr.................LeuSerCysProMe   536
     ||| ::|||||||||
1744 AACCAGCATTTTTCAAGCACGGCGCACAGCGGAACTCAAATGTTTCCA  1793

 536 tGluSerArgHisAlaThrTyrSerTrpArgHisLysGluAsnValGluG   553
     :::||||:: ||||:::||||||:::||:: :|||||:::::|||
1794 AAAGTCCAACTAGCCCGGGTGGTATGGAAGTTCCAGAATGGCGAGTTGA  1843

 553 lnSerCysGluPro.................GlyHisGlnSerProAsnCys   564
     :::::::|||:::||                  |||:::
1844 AGGCCCGCAAGTCCCAAGTACGGCTTTGTGGGGGAGGAGCAC.........  1884

 565 IleLeuPheIleGluAsnLeuThrAlaGlnGlnTyrGlyHisTyrPheCy   581
         :::||||:::     ||||| :::|||:::||:::|||||
1885 ..CTGCTCATCTTCAACCTGTCGGCACGGAGACAGCGGCGTGTACCAGTG  1931

 581 sGluAlaGlnGlyGluGlySerTyrPheArgGluAlaGlnHisTrpGlnLeuL   598
     :::::: ::::      |||:::                :::::||||||
1932 CCTGTCAGAGGAA.......AGGGTGAGGAATAAAACCGTCTCCCAGCTGC  1975

 598 euProGluAspGlyIleMetAlaGluHisLeuLeu   609
     ||                  :||:::||:::|||
1976 TG................GCCAAGCACGTTCTG  1992
```

```
 513 TGGCAAAGGAAGATGCCCCTTCGACCCGCCACACCTACACACAGTCA   562
     :::||| ::::||||||:::||           :::||
 150 heGluGlyAspGluValTyrSerThrIleArgLysGlnGluTyrAsn...   165
     |||:::|||  |||||||:::|||:::
 563 TGGTTGGGGCGAGCTCACTCT........GGGACGTCCTATAATTTC   603

 166 ...GlyLysIleProArgPheArgArgIleArgGlyIleArgGlyGluLeuTyrGluLeuuTy   181
        ||| ::|||  |||  :::|||     :::
 604 TTGGGCAGTGAACCCATCATCTCTGAAACTCTTCCACAGTCCCTTGAG   653

 181 rThrSerAspThrVal......MetGlnAsnProGlnPheIleLysAlaT   196
     :||:::|||  :::        |||:::|||:::|||
 654 GACGGAGTATGCCATCCGTGGCTGGCTGAACGAGCCTAGCTTCGTCTTGCTG   703

 196 hrIleValHisGlnAsp...............GlnAlaTyrAspAspLysIle   208
     :::::: ||||||           |||||||  :::|||
 704 ACGTGATCCAGAAAAGCCCAGATGTCCGGAGGGTGAAGATGACAAGGTC   753

 209 TyrTyrPheArgGluAspLysAsnProAlaLysAsnProGluAlaProLe   225
     |||:::|||  :: |||:::|||||| :: |||:::|||:::|||||
 754 TACTTCTTTTTTACGGAGGTATCCGTGGAGTACGGAATTCGTCTTCAAGTT   803

 225 uAsnValSerArgValAlaGlnLeuCysArgGlyLysAspGlnGlyGlyGluS   242
     :::|||:::|||  :::||||||:::||||||:::
 804 GATGATCCCGGCAGTTGCCAGGGTGCAAGGGCGACCAGGGCGGCCTGC   853

 242 erSerLeuSerValSerLysTrpAsnThrPheLeuLysAlaMetLeuuVal   258
     ::|||:::|||||||   |||:::|||||||||||:::
 854 GGACTTTGCAA...AAAAGTGGACCTCCTTCCTAAAGGCCAGGCTGATC   900

 259 CysSerAspAlaAlaThrAsnLysAsnPheAsnArgLeuGlnAspValPh   275
     ||||||  |||:::||||||||||||||||||
 901 TGCTCCAAGCCAGACAGTGGCCTGGTCTTCAACATACTTCAGGATGTGTT   950

 275 eLeuLeuProAspProSerGlyGlnTrpArgAspThrArgValTyrGlyV   292
     :|||||||||||:||:||:|||:::
 951 TGTGCTGAGGGCCCCGGGC.....CTCAAGGAGCCTGTGTTCTATGCGGG   994

 292 alPheSer........AsnProTrpAsnTyrSerAlaValCysValTyr   305
     |||||||          :|||:::|||:::|||:::
 995 TCTTCACCCCACAGCTGAACAATGTGGGTCTGTCAGCGGTGTGCGCCTAC  1044

 306 SerLeuGlyAspIleAspLysValPhe......................   314
     ||||||:::
1045 ACACTGGCCACGGTGGAGGCAGTCTTCTCCCGTGAAAGTACATGCAGAG  1094

 315 .......ArgThrSerSerLeuLys.....GlyTyrHisSerSerL   326
            :::||||||:::         |||:::
1095 TGCCACAGTGGAGCTCTCACACCAGTGGGTGCCTACAATGGCCCAG  1144

 326 euProAsnProArgProGlyLysCysLeuCysProAspGlnGln.......   339
     ||:::|||:::|||||||||||||:::
1145 TGCCCACTCCCCGACCTGGAGCCGTGTATCGACAGTGACACGTGAGGCCCGGGCAGCC  1194

 340 ...............ProIleProThrGluThrPheGlnValAlaAs   350
                    |||  :::|||
1195 AACTACACCAGCTCCTTGAATCTCCCAGACAAAACACTGCAGTTTGTAAA  1244

 350 pArgHisProGluValAlaGlnArgValGluProMetGlyProLeuLysT   367
     :|||||||:::      ::: :::
1245 AGACCACCCCTTGATGGATGACTCAGTGACCCGATAGACAAC..AGAC  1291

 367 hrProLeuPheHisSerLysTyrHisTyrGlnLysValAlaValHisArg   383
     :::||||||  :::|||:::
1292 CCAAGCTGATCAAAAAGATGTAAAACTACACCCAGATAGTGGTAGACAGG  1341

 384 MetGlnAlaSerHisGlyLeuThrPheHisValLeuTyrLeuThrThrAs   400
     |||:::|||   :::||||||:::|||:::||||||
1342 ACCCAGCGCCCTGGATGGGACTTTCTACGACGTCATGTTCATCAGCACAGA  1391

 400 pArgGlyThrIleHisLysValValGluProGlyGlyGlnHisSerP   417
     |||||||:::|||:::
```

Wed Jan 19 11:15:40 2000

OM of: US-09-041-236-2 to: Issued_Patents_NA:*   out_format : pfs

Date: Jan 19, 2000 4:14 AM

About: Results were produced by the GenCore software, version 4.5,
      Copyright (c) 1993-1998 Compugen Ltd.

Command line parameters:
-MODEL-frame+ p2n.model -DEV-xlp
-Q-/cgn2_1/USPTO_spool/US09041236/runat_15012000_164052_514/app_query.fasta.1
-DB-Issued_Patents_NA -QFMT-fastap -SUFFIX-LEE.rni
-GAPOP-12.000 -GAPEXT-4.000 -MINMATCH-0.100 -LOOPCL-0.000
-LOOPEXT-0.000 -QGAPOP-4.500 -XGAPOP-10.000
-XGAPEXT-0.500 -FGAPOP-6.000 -FGAPEXT-7.000 -YGAPOP-10.000
-YGAPEXT-0.500 -DELOP-6.000 -DELEXT-7.000 -START-1
-MATRIX-blosum62 -TRANS-human40.cd1 -LIST-45 -DOCALIGN-200
-THR_SCORE-pct -ALIGN-15 -MODE-LOCAL -OUTFMT-pfs -NORM-ext
-MINLEN-0 -MAXLEN-1000000 -USER-US09041236 -NCPU-6 -ICPU-3
-NO_XLPXY -WAIT -THREADS-1

Search information block:
Query: US-09-041-236-2
Query length: 634
Database: Issued_Patents_NA:*
Database sequences: 207703
Database length: 57918364
Search time (sec): 27.510000

score_list:

Sequence	Strd	Orig	ZScore	EScore	Len	! Documentation
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-121-713D-53 +	649.50	1221.80	1.2e-60	2601	!
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-835-268-53 +	649.50	1221.80	1.2e-60	2601	!
/cgn2_6/ptodata/2/ina/5D_COMB.seq:US-09-060-692-53 +	649.50	1221.80	1.2e-60	2601	!
/cgn2_6/ptodata/2/ina/PCTUS9_COMB.seq:PCT-US94-10151A-53 +	649.50	1221.80	1.2e-60	26	!
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-136-922-1 +	547.00	1028.95	6.9e-50	1481	!
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-121-713D-57 +	449.00	828.87	9.6e-39	2854	!
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-835-268-57 +	449.00	828.87	9.6e-39	2854	!
/cgn2_6/ptodata/2/ina/5D_COMB.seq:US-09-060-692-57 +	449.00	828.87	9.6e-39	2854	!
/cgn2_6/ptodata/2/ina/PCTUS9_COMB.seq:PCT-US94-10151A-57 +	449.00	828.87	9.6e-39	28	!
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-121-713D-63 +	408.00	750.49	2.2e-34	2504	!
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-835-268-63 +	408.00	750.49	2.2e-34	2504	!
/cgn2_6/ptodata/2/ina/5D_COMB.seq:US-09-060-692-63 +	408.00	750.49	2.2e-34	2504	!
/cgn2_6/ptodata/2/ina/PCTUS9_COMB.seq:PCT-US94-10151A-63 +	408.00	750.49	2.2e-34	25	!
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-121-713D-61 +	406.00	745.74	4.1e-34	2670	!
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-835-268-61 +	406.00	745.74	4.1e-34	2670	!
/cgn2_6/ptodata/2/ina/5D_COMB.seq:US-09-060-692-61 +	406.00	745.74	4.1e-34	2670	!
/cgn2_6/ptodata/2/ina/PCTUS9_COMB.seq:PCT-US94-10151A-61 +	406.00	745.74	4.1e-34	26	!
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-121-713D-59 +	399.00	728.29	3.8e-33	3560	!
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-835-268-59 +	399.00	728.29	3.8e-33	3560	!
/cgn2_6/ptodata/2/ina/5D_COMB.seq:US-09-060-692-59 +	399.00	728.29	3.8e-33	3560	!
/cgn2_6/ptodata/2/ina/PCTUS9_COMB.seq:PCT-US94-10151A-59 +	399.00	728.29	3.8e-33	35	!
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-121-713D-55 +	364.00	672.82	4.7e-30	1332	!
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-835-268-55 +	364.00	672.82	4.7e-30	1332	!
/cgn2_6/ptodata/2/ina/5D_COMB.seq:US-09-060-692-55 +	364.00	672.82	4.7e-30	1332	!
/cgn2_6/ptodata/2/ina/PCTUS9_COMB.seq:PCT-US94-10151A-55 +	364.00	672.82	4.7e-30	13	!
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-121-713D-65 +	117.50	208.11	0.0004	369	!
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-835-268-65 +	117.50	208.11	0.0004	369	!
/cgn2_6/ptodata/2/ina/5D_COMB.seq:US-09-060-692-65 +	117.50	208.11	0.0004	369	!
/cgn2_6/ptodata/2/ina/PCTUS9_COMB.seq:PCT-US94-10151A-65 +	117.50	208.11	0.0004	36	!
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-07-997-133-2 +	116.00	180.21	0.0129	2469	!
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-451-822A-14 +	116.00	179.22	0.0147	2662	!
/cgn2_6/ptodata/2/ina/5C_COMB.seq:US-07-997-133-2 +	116.00	180.21	0.0129	2469	!
/cgn2_6/ptodata/2/ina/5C_COMB.seq:US-08-451-822A-14 +	116.00	179.22	0.0147	2733	!
/cgn2_6/ptodata/2/ina/PCTUS9_COMB.seq:US-07-997-133-2 +	116.00	180.21	0.0129	2469	!
/cgn2_6/ptodata/2/ina/PCTUS9_COMB.seq:US-07-997-133-14 +	114.00	174.97	0.0253	2733	!
/cgn2_6/ptodata/2/ina/PCTUS9_COMB.seq:US-08-371-001-14 +	114.00	174.97	0.0253	273	!
/cgn2_6/ptodata/2/ina/PCTUS9_COMB.seq:US-08-459-296-1 +	109.00	166.53	0.0746	2469	!
/cgn2_6/ptodata/2/ina/3A_COMB.seq:US-08-168-917-1 +	106.00	150.33	0.5964	5427	!
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-460-510-1 +	106.00	150.33	0.5964	5427	!
/cgn2_6/ptodata/2/ina/5C_COMB.seq:US-08-460-490-1 +	106.00	150.33	0.5964	5427	!
/cgn2_6/ptodata/2/ina/PCTUS9_COMB.seq:US-09-007-30-1 +	106.00	150.33	0.5964	5427	!
/cgn2_6/ptodata/2/ina/PCTUS9_COMB.seq:US-92-00730-1 +	106.00	150.33	0.5964	5427	!
/cgn2_6/ptodata/2/ina/PCTUS9_COMB.seq:US-92-00862-1 +	106.00	150.33	0.5964	542	!
/cgn2_6/ptodata/2/ina/5D_COMB.seq:US-09-016-366A-24 -	104.00	167.35	0.0672	1103	!

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-121-713D-53

seq_documentation_block:
; Sequence 53, Application US/0812l713D
; Patent No. 5639856
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Matthes, David
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/121,713D
; FILING DATE: 13-SEP-1993
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36, 627
; REFERENCE/DOCKET NUMBER: B94-002-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)343-4341
; TELEFAX: (415) 343-4342
; TELEX:
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2601 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 16..2331
US-08-121-713D-53

alignment_scores:
     Quality: 649.50   Length: 687
     Ratio: 1.620   Gaps: 27
     Percent Similarity: 58.370   Percent Identity: 28.093

alignment_block:
US-09-041-236-2 x US-08-121-713D-53

Align seq 1/1 to: US-08-121-713D-53 from: 1 to: 2601

1 LeuLeuLeuLeuTrpAlaAlaAlaAlaSerAlaGlnGlyHisLeuAr 17
   :::::: :::|:|||:::|||::::|||||||::::::::::::
34 ATTGTCTGTCTTTCTCGGGAGTATTACTTACAGCAAGAGCAAACTATCA 83

17 gSerGly.........ProArgIlePheAlaAlaValTrpLysGlyHisV 30
   :::|::::   :::||||:::
84 GAATGGGAAGAACAATGCCCAAGGCTGAAATTATCCTACAAAGAAATGT 133

```
30  alGlyGlnAspArgValAsp..........PheGlyGlnThrGlu  41
    ::   :::::::||||:||
134 TGGAATCCAACAATGTGATCACTTTCAATGGCTTGGCCAACAGCTCCAGT  183

42  ProHisThrValLeuPheHisGluProGlySerSerValTrpValG1  58
    ||||||| ::::::||||| ||| ::::||||||
184 TATCATACCTTCCTTTTTGGATGAGGAA...CGGAGTAGGCTGTATGTTGG  230

58   yGlyArgGlyLysValTyrLeuPheAspPheProGluGlyLysAsnAlaS  75
    |:::::: || |||||||| ||| |||:::
231 AGCAAGGATCACATATTTCATTCGACCTGGTTAATATCAAGGATTTTC  280

75  ervalArgThrValAsnIleGlySerThrLys..........  85
    ::  :::::: |||::
281 AAAAGATTGTGTGGCCAGTATCTTACACCAGAAGAGATGAATGCAAGTGG  330

86  ..GlySerCysLeuAspLysArgAspCysGluAsnTyrIleThrLeuLe  101
    |||:|| ||||||| :::|||
331 GCTGGAAAAGACATCCTGAA...GAATGTGCTAATTTCATCAAGGTACT  377

101 uGluArgSerGlu...GlyLeuLeuAlaCysGlyThrAsnAlaArgH  117
    |::: :::::: || |||||||||:::|||
378 TAAGGCATATAATCAGACTCACTTGTACGCCCTGTGGAACGGCGGCTTTTC  427

117 1sProSerCysTrpAsnValAsnGlyThrValProLeuGlyGlu  133
    |||| ||| ||:::|||||:
428 ATCCAATTTGC..........ACCTACATTGAAATTGGACAT  459

134 Met..........  134
460 CATCCTGAGGACAATATTTTTAAGCTGGAGAACTCACATTTTGAAAACGG  509

135 .ArgGlyTyrAlaProPheSerProAspGluAsnSerLeuValLeuPheG  151
    ||||||||::: ::: ||:|||| |||::|:
510 CCGTGGGGAAGAGCCATATGACCCTAAGCTGCTACAGCATCCCTTTTAA  559

151 luGlyAspGluValTyrSerThrIleArgLysGlnGluTyrAsnGlyLys  167
    ||:::||||||||||:::||
560 TAGATGGAGAATTATACTCTGGAACTCGCAGCTGATTTTATGGGGCGAGAC  609

168 IleProArgPheArgArgIleArgGlyGlyGluSerGluTyrThr......  182
    |||||| ::: ||:
610 TTTGCTATCTTCCGAACTCTGGGCACCACCACCCAATCAGGACAGAGCA  659

183 ....SerAspThrValMetGlnAsnProGlnPheIleLysAlaThrIlev  198
    ::::::||||||::::::||||:||||||| :::
660 GCATGATTCCAGGTGGCTCAATGATCCAAAGTTCATTAGTGCCCACCTCA  709

198 alHisGln...AspGlnAlaTyrAspAspTyrIleTyrTyrPhePheArg  213
    ::   |||::: ||||||| |||||| ::::::
710 TCTCAGAGAGTGACAATCCTGAAGATGACAAAGTATACTTTTTCTTCCGT  759

214 GluAspAsnProAspLysAsnProGluAlaProLeuAsnValSerArgVa  230
    ||||||||::: ::::::|||::||
760 GAAAATGCAATAGATGGAGAATGATGGAAAAGCTACTCACGCTAGAAT  809

230 1AlaGlnLeuCysTyrArgGlyAspGlnGlyGlySerLeuSerValS  247
    :::::::||||||||||||||||||| |||
810 AGGTCAGATATGCAAGAATGACTTTGGAGGGCACAGAGTCTG...GTGA  856

247 erLysTrpAsnProSerGlyGlnTrpArgAspThrArgValTyrGlyV  263
    ::|||||||||| :::||||||||
857 ATAAATGGACAACATTCCTCAAAGCTCGTCTGATTTGCTACGGTGTAGAG  906

264 ThrAsn..........LysAsnPheAsnArgLeuGlnAsnValPheLeuGlu  277
    |||  ::||:||||:||||||||||
907 CCAAATGGCATTGACACTCATTTTGATGAACTGCAGAATGTATTCCTTAAT  956

277 u......ProAspProSerGlyGlnTrpArgAspThrArgValTyrGlyV  292
    ||:: ::: |||||||||
957 GAACTTTAAAGATCCT..........AAAAATCCAGTTGTATATGGAG  994

292 alPhe..........SerAsnProTrpAsnTyrSerAlaValCysValTyr  305
    ::: |||
```

```
995 TGTTTACGACTTCCAGTAACACATTTTCAAGGGATCAGCCGTGTGTATGTAT  1044

306 SerLeuGlyAspIleAspLysValAlaPhe..........  314
    ::||| :::|||||||
1045 AGCATGAGTGATGTGGAAGGGTGTTCCTTGGTCCCATATGCCACCAGGGA  1094

315 ....ArgThrSerLeuLysGlyTyrHisSerSerLeuProAsnProA  330
    ::| |||
1095 TGGACCCAACTATCAATGGCTGCCTTATCAAGGAAGAGTCCCCTATCCAC  1144

330 rgProGlyLysCys..........LeuProAspGlnGln  339
    |||||||| |||  ::: :::
1145 GGCCAGGAACTTGTCCCACAAAACATTTGGTGGTTTTGACTTCTACAAAG  1194

340 ProIleProThrPheGluThrPheGlnValAlaAspArgHisProGluValAl  356
    :::||| :::|||||||||||||:
1195 GACCTTCCTGATGATGTTATAAACCTTTGCAAGAAGTCATCCAGCCATGTA  1244

356 aGlnArgValGluProMetGly..........ProLeuLysThrProL  369
    |||  ||| ||||||||||:::
1245 CAATCCAGTGTTTCCTATGAACAATCGCCAATAGTGATCAAAACGGATG  1294

369 euPheHisSerLysTyrHisTyrGlnLysValAlaValHisArgMetGln  385
    ::  ::::::||:::
1295 TA........AATTATCAATTTACACAAATTGCTAGACCCAGGTGGAT  1335

386 AlaSerHisGlyThrPheHisValLeuTyrLeuThrThrAspArgGl  402
    ::::::::|||||||||| ||||| ||
1336 GCAGAAGATGGACAG...TATGATGTTATGTTTATCGGAACAGATGTTGG  1382

402 yThrIleHisLysValValGlu...ProGlyGluGlnGluHisSerPheA  418
    ||:::||  ::::::||| |||
1383 GACCGGTTCTTAAGTAGTTTCAATTCCTAAGGAGACTTGGTATGATTTAG  1432

418 laPheAsnIleMet...GluIleGlnProPheArgArgAlaAlaAlaIle  433
    :::::: |||  ::::::||||||
1433 AAGAGGGTTCTGCTGGAAGAAAATGACAGTTTTCGGGAACCGACTGCTATT  1482

434 GlnThrMetSerLeuAspAlaGluArgArgLysLeuTyrValSerSerGl  450
    |||||||| ||| ||||||||:::||:::
1483 TCAGCAATGGAGCTTTCCACTAAGCAGCAACAACTATATATTGGTTCAAC  1532

450 nTrpGluValSerGlnValProLeuAspLeuCysGluValTyrGlyGlyG  467
    ||||::::::|||||||||||:::||||||
1533 GGCTGGGGTTGCCAGCTCCCTTACACCGGTGTGATATTTACGGGGAAAG  1582

467, lyCysHisGlyCysLeuMetSerArgArgAspProTyrCysGlyTrpAspGln  483
    |:::|| :::||||||||||||||:::|||||
1583 CGTGTGCGTTGCCCGCCGCCCTTACTGTGCTTGGGATGGT  1632

484 GlyArgCysIleSerIleTyr..........SerGluArgSerValLeuGl  498
    ::: :::::: |||
1633 TCTGCATGTTCTCGCTATTTCCCACTGCAAAGAGACGCCACAAGACGACA  1682

498 nSerIleAsnProAlaGluProHisLysLysGluCysProAsnProLysProG  515
    |:::::||| :::::||
1683 AGATATAAGAAATGGAGACCCACTGACTCACTGTTCAGACTTACACCATG  1732

515 sPLys..........AlaProLeuGlnLysValSerLeu...AlaPro  526
    |||| :::
1733 ATAATCACCATGGCCACACCCCTGAGAGAGAATCATCTATGGTGTAGAG  1782

527 AsnSerArgTyrTyrLeuSerProMetGluSerArgHisAlaThrTy  543
    |||||| :::::::||:::
1783 AATAGTACACATTTTGGAATGCAGTCCGAGTCGCCAGAGAGCGGCTGGT  1832

543 rSerTrpArgHisLysGluAsnValGluGlnSerCysGlyValHisG  560
    |||  :::::: |||
1833 CTATTGGCAATTCCAGAGCGCAGGGCAAATGAAGAGCGAAAGAAGATCAGAG  1882

560 lnSerProAsnCysIle..........LeuPheIleGluAsnLeu  571
    ::: |||
```

1883 TGGATGATCATATCATCAGGACAGATCAAGGCCTTCTGCTACGTAGTCTA 1932

572 ThrAlaGlnGlnTyrGlyHisTyrPheCysGluAlaGlnGluGlySerTy 588
  ||||:||||||||||||||::||:|||||| ||| |||:::
1933 CAACGAAGGATTCAGGCAATTACCTCTGCCATGCGGGTGAACATGGGTT 1982

588 rPheArgGluAlaGlnHisTrpGlnLeuLeuProGluAspGlyIleMetA 605
  ::||:|| ||| |||::: ||| |||
1983 CATACAAACTCTTTAAGGTAACCCTG........GAAGTCATTGACA 2023

605 laGluHisLeu 608
  ::||||||||||
2024 CAGAGCATTTG 2034

; seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-835-268-53

; seq_documentation_block:
; Sequence 53, Application US/08835268
; Patent No. 5807826
; GENERAL INFORMATION:
;   APPLICANT: Goodman, Corey S.
;   APPLICANT: Kolodkin, Alex L.
;   APPLICANT: Matthes, David
;   APPLICANT: Bentley, David R.
;   APPLICANT: O'Connor, Timothy
;   TITLE OF INVENTION: The Semaphorin Gene Family
;   NUMBER OF SEQUENCES: 100
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
;   STREET: 268 Bush Street, Suite 3200
;   CITY: San Francisco
;   STATE: CA
;   COUNTRY: USA
;   ZIP: 94104
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/835,268
;   FILING DATE:
;   CLASSIFICATION:
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US/08/121,713
;   FILING DATE: 13-SEP-1993
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Osman, Richard A.
;   REGISTRATION NUMBER: 36,627
;   REFERENCE/DOCKET NUMBER: B94-002-1
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (415)343-4341
;   TELEFAX: (415) 343-4342
;   TELEX:
; INFORMATION FOR SEQ ID NO: 53:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 2601 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 16..2331
US-08-835-268-53

alignment_scores:
  Quality: 649.50   Length: 687
  Ratio: 1.620   Gaps: 27
Percent Similarity: 58.370 Percent Identity: 28.093

alignment_block:

US-09-041-236-2 x US-08-835-268-53

Align seg 1/1 to: US-08-835-268-53 from: 1 to: 2601

1 LeuLeuLeuLeuTrpAlaAlaAlaAlaSerAlaGlnGlyHisLeuAr 17
 :::: ||:||:||::|| |||||||| |||| :: ::::
34 ATTGTCTGTCTTTTCTGGGGAGTATTACTTACAGCAAGAGCAAACTATCA 83

17 gSerGly........ProArgIlePheAlaValTrpLysGlyHisV 30
 ::::|| :::||
84 GAATGGGAGAACAATGTGCCAGGCTGAAATTATCCTACAAAGAAATGT 133

30 alGlyGlnAspArgValAsp.............PheGlyGlnThrGlu 41
 ::||||||||||||
134 TGGAATCCAACAATGTGATCACTTTCAATGCTTGGCCAACAGCTCCAGT 183

42 ProHisThrValLeuPheHisGluProGlySerSerSerValTrpValGl 58
 |||||| ||:|| || ||| ::::||:||||
184 TATCATACCTCCTTTGGATGAGGAA...CGGAGTAGGCGTGTATGTTGG 230

58 yGlyArgGlyLysValTyrLeuPheAspPheProGluGlyLysAsnAlaS 75
 |:::||||||||:||||||| |::||| ||||||||
231 AGCAAGGATCACATATTTCATTCGACCTGGTTAATATCAAGGATTTTC 280

75 erValArgThrValAsnIleGlySerThrLys.............. 85
 ::: |||||||||||||||
281 AAAGATTGTGTGGCCAGTATCTTCACCAGAAGAGATGAATGCAAGTGG 330

86 ...GlySerCysLeuAspLysArgAspCysGluAsnTyrIleThrLeuLe 101
 |||: || ||||| |||::||| ::||
331 GCTGGAAAAGACATCCTGAAA...GAATGTGCTAATTTCATCAAGGTACT 377

101 uGluArgArgSerGlu...GlyLeuLeuAlaCysGlyThrAsnAlaArgH 117
 |:::|| ::: |||||||||||
378 TAAGGCATATAATCAGACTCACTTGTACGCCTGTGGAACGGGGCTTTC 427

117 isProSerCysTrpAsnLeuValAsnGlyThrValValProLeuUGlyGlu 133
 |||| || :: :|||||
428 ATCCAATTGC.........ACCTACATTGAAATTGGACAT 459

134 Met.. 134
460 CATCCTGAGGACAATATTTTAAGCTGGAGAACTCACATTTTGAAAACGG 509

135 .ArgGlyTyrAlaProPheSerProAspGluAsnSerLeuValLeuPheG 151
 |||||| :::||:||:|| |||||
510 CCGTGGAAGAGTCCATATGACCCTAAGCTGCTGACAGCATCCCTTTAA 559

151 luGlyAspGluValTyrSerThrIleArgLysGlnGluTyrAsnGlyLys 167
 |||:|||||| ::: :||||||::|||||||||||
560 TAGAATGCAGAATTATACTCTGGAACTGCAGCTGATTTATGGGGCGAGAC 609

168 IleProArgPheArgIleArgGlyIleGluSerGluLeuTyrThr...... 182
 |||||| ::: ::: :||:
610 TTTGCTACTTCCGAACTCTTGGGCACCACCCAATCAGGACAGAGCA 659

183 ...SerAspThrValMetGlnAsnProGlnPheLeuLysAlaThrIleV 198
 :::||||||:::||:|||||| |||||
660 GCATGATTCCAGGTGCTCAATGATCCAAAGTTCATTAGTGCCCACCTCA 709

198 alHisGln...AspGlnAlaTyrAspAspLysIleTyrTyrPhePheArg 213
 :: ||:|| ::||||||:||||:||:|||||||||||
710 TCTCAGAGAGTGACAATCCTGAAGATGACAAAGTATACTTTTCTTCCGT 759

214 GluAspAsnProAspLysAsnProGluAlaProLeuAsnValSerArgVa 230
 |||:|| ::: ::: :::
760 GAAAATGCAATAGATGGAGAACCTCGAAAAGTACTCACGCTAGAAT 809

230 lAlaGlnLeuCysArgGlyAspGlnGlyGlyGlySerSerLeuServalS 247
 :::||||::||||:::|| ||||||| |||:
810 AGGTCAGATATGCAAGAATGACTTTGGAGGGCACAGAAGTCTG...GTGA 856

247 erLysTrpAsnThrPheLeuValAlaMetLeuValCysSerAspAlaAla 263

```
857 ::|||||::::|||||||||| ||||::||||         ::::     906
    ATAAATGGACAACATTCCTCAAAGCTCGTCTGATTTGCTCAGTGCCAGGT

264 ThrAsn........LysAsnPheAsnArgLeuGlnAspValPheLeuLeu  277
    |||||          :::||||||||||||||||||||
907 CCAAATGGCATTGACACTCATTTTGATGAACTGCAGGATGATTCCTAAT    956

277 u......ProAspProSerGlyGlnTrpArgAspThrArgValTyrGlyValV  292
    ||||||               ::::::  ||||||||||||||
957 GAACTTTAAAGATCCT..........AAAAATCCAGTTGTATATGGAG      994

292 alPhe........SerAsnProTrpAsnTyrSerAlaValCysValTyr     305
    |||||         |||||  :::::  ||||||||||||::||
995 TGTTTACGACTTCCAGTAACATTTCAAGGACTCAGCCGTGTGTATGTAT   1044

306 SerLeuGlyAspIleAspLysValPhe...................        314
    ||||||||||||||||||||
1045 AGCATGGAGTGATGTGAGAAGGGTGTTCCTTGGTCCATATGCCCACAGGGA 1094

315 ....ArgThrSerLeuLysGlyTyrHisSerSerLeuProAsnProA       330
       :::    |||||||||||  ::::  |||||  ||||
1095 TGGACCCAACTATCAATGGGTGCCTTATCAAGGAAGAGTCCCCTATCCAC  1144

330 rgProGlyLysCys................LeuProAspGlnGln        339
    ||||||||| |||                 :: :: :::
1145 GGCCAGGAACTTGTCCCCAGCAAAACATTTGGTGGTTTTGACTCTACAAAG 1194

340 ProIleProThrGluProPheGlnValAlaAspArgHisProGluValAl   356
    |||                      ||| ||||||||||||
1195 GACCTTCCTGATGATGTTATAACCTTTGCAAGAGAGTCATCCAGCCATGTA 1244

356 aGlnArgValGlyUProMetGly.........ProLeuLysThrProL     369
    :::  |||||||||||||||                :::||||
1245 CAATCCAGTGTTTCCTATGAACAATCGCCAATAGTGATCAAAACGGATG   1294

369 euPheHisSerLysTyrHisTyrGlnLysValAlaValHisArgMetGln  385
    ::  ::||||:::::|||||||||  |||  |||||
1295 TA......AATTATCAATTTACACAAATGTCGTAGACCGAGTGGAT      1335

386 AlaSerHisGlyGluThrPheHisValLeuTyrLeuThrThrAspArgGl   402
    |||||   |||||||||::  ||| :::||  ||||| ||
1336 GCCAAGAGTGGACAG...TATGATGTTATGTTTATCGGAACAGATGTTGG  1382

402 yThrIleHisLysValValGlu...ProGlyGluGlnGluHisSerPheA  418
    ||||||||||||||||||                 ::::::
1383 GACCGTTCTTAAAGTAGTTTCAATTCCTAAGGAGACTTGGGATGATTTAG 1432

418 laPheAsnIleMet...GluIleGlnProPheArgArgAlaAlaAlaIle  433
    :::::|||||||||      |||||||:::  ::|||||||
1433 AAGAGGTTCTGCTGGAAGAAATGACAGTTTTCGGGAACCGACTGCTATT  1482

434 GlnThrMetSerLeuAspGluArgAlaGluArgLysLeuTyrValSerSerGl 450
    |||||||||||:::  |::||||| ||||  |||||||||:::
1483 TCAGCAATGGGAGCTTTCCACTAAGCAGCAACAACTATATATTGGTTCAAC  1532

450 nTrpGluValSerGlnValProLeuLeuAspLeuCysGluValTyrGlyGlyG 467
    :::    |||||||||||||||||:::::|||  ||||
1533 GGCTGGGGTTGCCCAGCTCCCTTACACCGGTGTGATATTTACGGGGAAAG  1582

467 lyCysHisGlyGlyCysLeuMetSerArgAspProTyrCysGlyTrpAspGln 483
    |:::::|||||||  :::  ||||||||||||||:::
1583 CGTGTGCTGAGTGGTTGCCTCGCCCGAGACCCCTACTGTGGTTGGGATGGT  1632

484 GlyArgCysIleSerIleTyr......SerSerGluArgSerValLeuGl   498
    :::  ||||||||||::::     :::  |||||||||
1633 TCTCGCATGTTCTCGCTATTTTCCCACTGCAAAGAGACGCCACAAGACGACA 1682

498 nSerIleAsnProAlaGluProHisLysGlyGluCysProAsnProLysPro  515
    |::::||||| |||||||  :: :: ||||||||
1683 AGATATAAGAAATGGAGACCCACCACTCACTGTTCAGACTTACACCATG   1732

515 sPLys..........AlaProLeuGlnInLysValSerLeu..AlaPro    526
    ||::                 :::::::
```

```
1733 ATAATCACCATGGCCAGCCCTGAAGAGAGAGATCATCATGTGTAGAG  1782

527 AsnSerArgTyrTyrLeuSerCysProMetGluSerArgHisAlaThrTy  543
    ||||||                 ::|||::::::|||
1783 AATAGTAGCACACATTTTGGAATGCAGTCCAGATCGCCAGAGACGGCTGGT 1832

543 rSerTrpArgHisLysGluAsnValGluGlnSerCysGlyProGlyHisG  560
    |::::                  ::::::  ||||  |||||
1833 CTATTGGCAATTCCAGAGGCGAAATGAAGAGCGAAAAGAAGAGATCAGAG  1882

560 lnSerProAsnCysIle.............LeuPheIleGluAsnLeu    571
    :::                          |||||||||:::
1883 TGGATGATCATATCATCAGGACAGACATCAAGGCCCTTCTGCTACGTAGTCTA 1932

572 ThrAlaGlnGlnTyrHisTyrPheCysGlnAlaGlnGluGlySerTy    588
    |||::::                   :::  ||||
1933 CAACAGAAGGATCAGGCAATTACCTCTGCCATGCGGTGGAACATGGGTT   1982

588 rPheArgGluAlaGlnHisTrpGlnLeuProGluAspGlyIleMetA    605
    ||::::                       :::  |||
1983 CATACAAACTCTTCTTAAGGTAACCCTG..........GAAGTCATTGACA 2023

605 laGluHisLeu 608
    ::|||||||||
2024 CAGAGCATTTG 2034
```

seq_name: /cgn2_6/ptodata/2/ina/5D_COMB.seq:US-09-060-692-53

seq_documentation_block:
; Sequence 53, Application US/09060692
; Patent No. 5935865
; GENERAL INFORMATION:
;   APPLICANT: Goodman, Corey S.
;   APPLICANT: Kolodkin, Alex L.
;   APPLICANT: Matthes, David
;   APPLICANT: Bentley, David R.
;   APPLICANT: O'Connor, Timothy
;   TITLE OF INVENTION: The Semaphorin Gene Family
;   NUMBER OF SEQUENCES: 100
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
;   STREET: 268 Bush Street, Suite 3200
;   CITY: San Francisco
;   STATE: CA
;   COUNTRY: USA
;   ZIP: 94104
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/060,692
;   FILING DATE:
;   CLASSIFICATION: 514
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US/08/121,713
;   FILING DATE: 13-SEP-1993
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Osman, Richard A.
;   REGISTRATION NUMBER: 36,627
;   REFERENCE/DOCKET NUMBER: B94-002-1
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (415)343-4341
;   TELEFAX: (415) 343-4342
;   TELEX:
; INFORMATION FOR SEQ ID NO: 53:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 2601 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
;   MOLECULE TYPE: cDNA

```
;  FEATURE:
;  NAME/KEY:    CDS
;  LOCATION:    16..2331
US-09-060-692-53

alignment_scores:
       Quality: 649.50          Length:       687
         Ratio:   1.620           Gaps:        27
Percent Similarity: 58.370    Percent Identity: 28.093

alignment_block:
US-09-041-236-2 x US-09-060-692-53

Align seg 1/1   to: US-09-060-692-53    from: 1   to: 2601
```

```
   1 LeuLeuLeuLeuTrpAlaAlaAlaAlaSerAlaGlnGlyHisLeuAr  17                    IIIIII:::::::    IIIIIIIII   710 GluAspAsnProAspLysAsnProGluAlaProLeuAsnValSerArgVa 230
     ::::::  III::III:::III:::::::    ::  IIIIIII                              :::  III::::IIIIIIIIII::::IIIIIIIIIII:::
  34 ATTGTCTGTCTTTCGGGGAGTATTACTTACAGCAAGAGCAAACTATCA  83                     710 TCTCAGAGAGTGACAATCCTGAGATCCTGAGACGACAAAGTATACTTTTCTTCCGT 759

  17 gSerGly..........ProArgIlePheAlaValTrpLysGlyHisV  30                     214 GluAspAsnProAspLysAsnProGluAlaProLeuAsnValSerArgVa 230
     IIIIII           IIIIIIII::::III::III                                    III::::IIIIIIIIII::::IIIIIIIIIII:::
  84 GAATGGAAGAACAATGTGCCAAGGCTGAAATTATCCTACAAGAAAATGT 133                    760 GAAAATGCAATAGATGGAGAACACTCTGGAAAGCTACTCACGCTAGAAT 809

  30 alGlyGlnAspArgValAsp..........PheGlyGlnThrGlu  41                        230 lAlaGlnLeuCysArgGlyArgGlyGlnGlyGlyGluSerLeuSerVals 247
     ::  IIIIIIII:III                                                         :::III:::III:::III    IIIIII    III:::   III
 134 TGGAATCCACAATGTGATCACTTTCAATGGCTTGGCCAACAGCTCCAGT 183                    810 AGGTCAGATATGCAAGATGAAGAATGACTTTGGAGGGCCACAGAAGTCTG...GTGA 856

  42 ProHisThrValLeuPheHisGluProGlySerSerValTrpValGl  58                      247 erLysTrpAsnThrPheLeuLysAlaMetLeuValCysSerAspAlaAla 263
     IIIIIII   III::IIII   III    III  :::::III                              ::III::III   III::IIIIIIII:::::::::::
 184 TATCATACCTTCCTTTGGATGAGGAA...CGGAGTAGGCTGTATGTTGG 230                    857 ATAAATGACAACATTCCTCAAAGCTCGTCTGATTGCTCAGTGCCAGT 906

  58 yGlyArgGlyLysValTyrLeuPheAspProGluProGlyLysAsnAlaS  75                    264 ThrAsn........LysAsnPheAsnArgLeuGlnAspValPheLeuLeu 277
     IIIIIIIIII   :::::III::::III:::IIII                                     III         :::::::III    IIIIIIIIIIII::
 231 AGCAAAGGATCACATATTTTCATTCGACCTGTTAATATCAAGGATTTTC 280                    907 CCAAATGGCAATTGACACTCATTTGATGAACTGCAGGATGATTCCTAAT 956

  75 ervalArgThrValAsnIleGlySerThrLys.............  85                        277 u.....ProAspProSerGlyGlnTrpThrArgValTyrGlyV 292
     IIIIIIIIIIIIIIII:::::III::    III                                        :                   ::::::      IIIIIIIII
 281 AAAAGATTGTGTGGCCAGATCTTACCACGAAGAGATGAATGCAAGTGG 330                     957 GAACTTTAAAGATCCT........AAAAATCCAGTTGTATATGGAG 994

  86 ...GlySerCysLeuAspLysArgAspCysGluAsnTyrIleThrLeuLe 101                   292 alPhe.........SerAsnProTrpAsnTyrSerAlaValCysValTyr 305
     III::::    ::  III  III    IIIIIII    ::III                             IIIII            IIIIII    III:::IIIIII::III
 331 GCTGGAAAGACATCCTGAA...GAATGTGCTAATTCATCAAGGTACT 377                      995 TGTTTACGCACTTCCAGTACACATTTCAAGGGATCAGCCGTGTGTATGTAT 1044

 101 uGluArgArgSerGlu...GlyLeuLeuAlaCysGlyThrAsnAlaArgH 117                   306 SerLeuGlyAspIleAspLeuLysValPhe................  314
     ::    IIIIIIIIIIII   IIIIIIIIIIII::::::III                              IIIIIIII:::   :::::IIIIIII
 378 TAAGGCATATATATCAGACTGCTCACTTGTACGCCTGTGAACGGGGCTTTTC 427                 1045 AGCATGAGTGATGTGAAGGGTGTTCCTTGGTCCATATGCCCACGGGA 1094

 117 lsProSerCysTrpAsnLeuValAsnGlyThrValProLeuGlyGlu 133                      315 ....ArgThrSerLeuLysGlyTyrHisSerSerLeuProAsnProA 330
     IIIII        III        III    :::::IIIIIIIIII                                :::      III    IIIIIIIII    ::::III
 428 ATCCAATTTGC...............ACCTACATTGAAATTGGACAT 459                      1095 TGGACCCAACTATCAATGGGTGCCTTATCAAGGAAGAGTCCCCTATCCAC 1144

 134 Met...............................  134                                  330 rgProGlyLysCys...............LeuProAspGlnGln 339
     IIIII                                                                    IIIIIIII::: III                ::: :::III
 460 CATCCTGAGGACAATATATTTTAAGCTGGAGAACTCACATTTTGAAAACGG 509                  1145 GGGCAGGGACTTGTCCCAGCAAAACATTGGTGTTTTGACTCTACAAAG 1194

 135 .ArgGlyTyrAlaProPheSerProAspGluAsnSerLeuValLeuPheG 151                   340 ProIleProThrPhGluThrPheGlnValAlaAspArgHisProGluValAl 356
     IIIIII :::::IIIIII::::III                    IIIII                       IIIIIIII   III     ::::III::::III    ::
 510 CCGTGGGAAGAGTCCCATATGACCCCTAAGCTGCTGACAGCATCCCTTTTAA 559                 1195 GACCTTCCTGATGATGTTATAACCTTTGCAAGAAGTCATCCAGCCATGTA 1244

 151 luGlyAspGluValTyrSerThrIleArgLysGlnGluTyrAsnGlyLys 167                   356 aGlnArgValGluProMetGly.........ProLeuLysThrProL 369
     IIIIIIIIIIIIIIIIIIIIIIIII      :::                                      :::   III  IIIIIIII    IIIIIIII    III
 560 TAGAATGGAGAATTATACTCTGGAACTGCAGCTGATTTTATGGGCGGAGAC 609                  1245 CAATCCAGTGTTTCCTATGAACAATCGCCCAATAGTGATCAAAACGGATG 1294

 168 IleProArgPheArgGlyIleArgGlyGlyGluTyrThr......  182                       369 euPheHisSerLysTyrHisTyrGlnLysValAlaValHisArgMetGln 385
     IIIIII ::::::III::::III                 III                             ::  :::::::III::::IIIIII    III    III:::
 610 TTTGCTATCTTCCGAACTCTTGGGCACCACCCAATCAGGACGACGAGCA 659                    1295 TA.......AATTATCAATTTACACAAATTGTCTAGACCCAGTGGAT 1335

 183 ...SerAspThrValMetGlnAsnProGlnPheLeuLysAlaThrIleV 198                    386 AlaSerHisGlyGluThrPheHisValLeuThrTyrLeuThrThrAspArgGl 402
     :::   IIIIIIII::::::III:::IIIIIIII::::III                                IIIIII:::  :::   III          III    III  II
 660 GCATGATTCCAGGTGGCTCAATGATCCAAGTTCATTAGTGCCCACCTCA 709                    1336 GCAGAAGATGGACAG...TATGATGTTATGTTTATCGGAACAGATGTTGG 1382

 198 alHisGln...AspGlnAlaTyrAspAspLysIleTyrTyrPhePheArg 213                   402 yThrIleHisLysValGlu..ProGlyGluGlnGlnHisSerPheA 418
     ::III                   IIIIIIIIIIIIIIIIIIIII:::                         IIIII IIIIIIIIIII    III       :::::::::: :::
                                                                             1383 GACCGTTCTTAAAGTAGTTTCAATTCCTAAGGAGACTTGGTATGATTTAG 1432

                                                                             418 laPheAsnIleMet...GluIleGlnProPheArgArgAlaAlaAlaIle 433
                                                                             :::IIIII         III:::    :::IIII    ::IIIIIII
                                                                             1433 AAGAGGTTCTGCTGGAGAAAATGACAGTTTTTCGGGAACCGACTGCTATT 1482

                                                                             434 GlnThrMetSerLeuAspAlaGluAlaArgLysLeuTyrValSerSerGl 450
                                                                             :::   III:::III::::III                 IIIIIIII
                                                                             1483 TCAGCAATGGAGCTTTCCACTAGCAGCAACAACTATATTGGTTCAAC 1532

                                                                             450 nTrpGluValSerGlnValProLeuAspLeuCysGluValTyrGlyGlyG 467
                                                                             III::III:::IIIIII:::IIII             IIIIIIIIIII
                                                                             1533 GGCTGGGGTTGCCCAGCTCCCTTTACACCGGTGTGATATTACGGGAAAG 1582

                                                                             467 lyCysHisGlyCysLeuMetSerArgAspProTyrCysGlyTrpAspGln 483
                                                                             :::III   IIIIIIIIIIIIIIIIIIIIIIIIIII:::
```

```
                                                    TYPE:  nucleic acid
                                                    STRANDEDNESS:  double
                                                    TOPOLOGY:  linear
                                                    MOLECULE TYPE:  cDNA
                                                    FEATURE:
                                                      NAME/KEY:  CDS
                                                      LOCATION:  16..2331
                                                    PCT-US94-10151A-53

                                            alignment_scores:
                                                    Quality:    649.50          Length:      687
                                                    Ratio:      1.620           Gaps:         27
                                            Percent Similarity:  58.370    Percent Identity:  28.093

                                            alignment_block:
                                            US-09-041-236-2 x PCT-US94-10151A-53

                                            Align seg 1/1  to:  PCT-US94-10151A-53   from: 1  to: 2601

1583  CGTGCGCTGAGTGTTGCCTCGCCCCAGACCCCTACTGTGCTTGGGATGGT 1632       1  LeuLeuLeuLeuTrpAlaAlaAlaAlaSerAlaGlnGlyHisLeuAr  17
                                                                      :::::  |||:||::||  :::|||  :::|||||:||||||||::
484  GlyArgCysIleSerIleTyr......SerSerGluArgSerValLeuGl  498        34  ATTGTCTGCTTTTCTCGGGAGTATTACTTACACGCAAGCAAAACTATCA  83
        :::  |||           ||||||:||::||:  ||
1633  TCTGCATGTGTCTCGCTATTTCCCACTGCAAAGACACGCCACAAGACGACA 1682      17  gSerGly..........ProArgIlePheAlaValTrpLysGlyHisV  30
                                                                      :::|||           |||||:|:        ::||||||
498  nSerIleAsnProAlaGluProHisLysGluCysProAsnProLysProA 515        84  GAATGGGAAGAACAATGTGCCAAGGCTGAAATTATCCTACAAAGAAATGT 133
        |::||:||::  :::||||  ::|::||||
1683  AGATATAAGAAATGGAGACCCACTGACTCACTGTTCAGACTTACACCATG 1732       20  aIGIyGlnAspArgValAsp..........PheGlyGlnThrGluG  41
                                                                      :|:                              :||
515  spLys..........AlaProLeuGlnLysValSerLeu...AlaPro  526         134  TGGAATCCAACAATGTGATCACTTTCAATGGCTTGGCCAACAGCTCCAGT 183
        |::|||        ::|:::||                ::||||
1733  ATAATCACCATGGCCACAGCCCTGAAGAGAATCATCTATGGTGTGTAGAG 1782      42  ProHisThrValLeuPheHisGluProGlySerSerValTrpValGl  58
                                                                      |||||  |||::  |||  :|:|||  :::::|||||||||||
527  AsnSerArgTyrTyrLeuSerCysProMetGluSerArgHisAlaThrTy 543        184  TATCATACCTTCCTTTGGATGAGGAA...CGGAGTAGGCTGTATGTTGG 230
        |||||      :::|||:::||||||           :::|||||||||
1783  AATAGTAGCACACATTTTGGAATGCAGTCCGAGTGGAAGAGCGCTGGT 1832        58  yGlyArgGlyLysValTyrLeuPheAspPheProGluGlyLysAsnAlaS  75
                                                                      |::|:  |||:||||:|  ||||:  |||||||||
543  rSerTrpArgHisLysGluAsnValGluGlnSerCysGluProGlyHisG 560        231  AGCAAAGGATCACATATTTTCATTCGACCTGGTTAATATCAAGGATTTC 280
        |||:  :::||        ||
1833  CTATTGGCAATTCCAGAGGCGAAATGAAGAGGCGAAAAGAAGAGGATCAGAG 1882     75  erValArgThrValAsnIleGlySerThrLys.................  85
                                                                      |||::
560  lnSerProAsnCysIle..........LeuPheIleGluAsnLeu  571          281  AAAAGATTGTGTGGCCAGTATCTTACACCAGAAGAGAATGAATGCAAGTGG 330
        ::::|                      |||||||||||:|:||||
1883  TGGATGATCATATCATCAGGACAGCAGATCAAGGCCTTCTGCTACGTAGTCTA 1932    86  ...GlySerCysLeuAspLysArgAspCysGluAsnTyrIlePheThrLeuLe 101
        ||||                        ::|||  :::||  :::||
572  ThrAlaGlnGlnTyrGlyHisIleTyrPheCysGluAlaGlnGluGlySerTy 588     331  GCTGGAAAAGACATCCTGAAA...GAATGTGCTAATTTCATCAAGGTACT 377
        |||:::||::||:  ::::|:::||||  |||  :|::::
1933  CAACAGAAGGATTCAGGCAATTACCCTGCCCATGCCGGTGGAACATGGGTT 1982     101  uGluArgArgSerGlu...GlyLeuLeuAlaCysGlyThrAsnAlaArgH 117
        |||::||  :::|||:::|:::|||                 |||
588  rPheArgGluAlaGlnHisTrpGlnLeuLeuProGluAspGlyIleMetA 605       378  TAAGGCATATAATCAGACTCACTTGTACGCCTGTGGAACGGGGGCTTTC 427
        :::|:::|||                :||
1983  CATACAAACTCTTCTTAAGGTAACCCTG........GAAGTCATTGACA 2023       117  lsProSerCysTrpAsnLeuValAsnGlyThrValTyrAlaProLeuGlyGlu 133
        :::::|||||::|:|||||                 |||        ||:||::|||
605  laGluHisLeu 608                                               428  ATCCAATTGC.................ACCTACATTGAAATTGGACAT 459
        ::||||||||                                                    |||:::
2024  CAGAGCATTG 2034                                              134  Met.........................................  134
                                                                      |||:::
seq_name: /cgn2_6/ptodata/2/1na/PCTUS9_COMB.seq:PCT-US94-10151A-53     460  CATCCTGAGGACACAATATTTTAAGCTGGAGAACTCACATTTGAAAACG 509
                                                                      :::|:::||  :::||  ||:::|||||::::||
seq_documentation_block:                                             135  .ArgGlyTyrAlaProPheSerProAspSerLeuValLeuPheG 151
;   Sequence 53, Application PC/TUS9410151A                             |||:::  ||  :::::||||:::|||||||||||
;   GENERAL INFORMATION:                                              510  CCGTGGGAAGAGTCCCATATGACCCTAAGCTGCTGACAGCCATCCCTTTAA 559
;       APPLICANT:   The Regents of the University of California       151  luGlyAspGluValTyrSerThrIleArgLysGlnGluTyrArgAsnGlyLys 167
;       TITLE OF INVENTION:  The Semaphorin Gene Family                   :::|||:::|||||||||||||||                 :::
;       NUMBER OF SEQUENCES:  66                                      560  TAGATGGAGAATATACCTGGAACTGCAGTGATTTATGGGGCGAGAC 609
;       CORRESPONDENCE ADDRESS:                                       168  IleProArgPheArgIleArgGlyGluSerGluGluTyrThr...... 182
;         ADDRESSEE:   FLEHR HOHBACH TEST ALBRITTON & HERBERT             :::  |:|||  :::|||||  :||  |||
;         STREET:   4 Embarcadero Center, Suite 3400                  610  TTTGCTATCTTCCGAACTTCTGGGCACCACCCAATCAGGACGAGAGCA 659
;         CITY:   San Francisco                                       183  ....SerAspThrValMetGlnAsnProGlnPheIleLysAlaThrIleV 198
;         STATE:   CA
;         COUNTRY:   USA
;         ZIP:   94111-4187
;   COMPUTER READABLE FORM:
;       MEDIUM TYPE:   Floppy disk
;       COMPUTER:   IBM PC compatible
;       OPERATING SYSTEM:   PC-DOS/MS-DOS
;       SOFTWARE:   PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:   PCT/US94/10151A
;       FILING DATE:   13-SEP-1994
;       CLASSIFICATION:
;   ATTORNEY/AGENT INFORMATION:
;       NAME:   Osman, Richard A.
;       REGISTRATION NUMBER:   36,627
;       REFERENCE/DOCKET NUMBER:   FP-58750-PC/RAO
;   TELECOMMUNICATION INFORMATION:
;       TELEPHONE:   (415) 781-1989
;       TELEFAX:   (415) 398-3249
;       TELEX:   910 277299 FHT UR
;   INFORMATION FOR SEQ ID NO:  53:
;       SEQUENCE CHARACTERISTICS:
;           LENGTH:   2601 base pairs
```

```
1533 GGCTGGGGTGCCAGCTCCCTTTACACCGGTGTGATATTTACGGGAAAG 1582
                :::::::::::::::::::::||||||::::|||
660 GCATGATTCCAGGTGGCTCAATGATCCAAAGTTCATTAGTGCCACCTCA 709

467 lyCysHisGlyCysLeuMetSerArgAspProTyrCysGlyTrpAspGln 483
            ||:|||         ||||:||||:::::||||:::|||||:
198 alHisGln...AspGlnAlaTyrAspLysIleTyrTyrPhePheArg 213

1583 CGTGTGCTGAGTGTGTCCTCCGCCCCGAGACCCTTACGTGCTGGGATGGT 1632
         ::  |||       ||||:||||||::::||||||:
710 TCTCAGAGAGTGACAATCCTGAACGATGACAAAGTATACTTTTTCTTCCGT 759

484 GlyArgCysIleSerIleTyr......SerSerGluArgSerValLeuGl 498
     |||  :::   :::::            ::  :::|||||
214 GluAspAsnProAspLysAsnProGluAlaProLeuAsnValSerArgVa 230

1633 TCTGCATGTTCTCGCTATTTTCCCACTGCAAAGACGCACAAGACGACACA 1682
     |||:|||      ::: :::||||    ::::|||||:
760 GAAAATGCAATAGATGGAGAACACTCTGAAAAGCTACTCACGCTAGAAT 809

498 nSerIleAsnProAlaGluProHisLysGluCysProLysProA 515
     |:::||||:    |||::::||||    ::||||
810 AGGTCAGATATGCAAGAATGACTTTGGAGGGCACAGAGAAGTCTG...GTGA 856

515 spLys..........AlaProLeuGlnLysValSerLeu...AlaPro 526
     :|||               :::||
857 ATAAATGGACAACATTCCTCAAAGCTCGTCTGATTTGCTCAGTGCCAGGT 906

1733 ATAATCACCATGGCCACAGCCCTGAAGAGAGAAATCATCTATGGTGTAGAG 1782
            |||||   :::|||||||||||:::||||||:
247 erLysTrpAspAsnThrPheLeuLysAlaMetLeuValCysSerAspAlaAla 263

527 AsnSerArgTyrTyrLeuSerCysProMetGluSerArgHisAlaThrTy 543
     |||||         ::::||||||||:||:||||||||||:::
264 ThrAsn.......LysAsnPheAsnArgLeuGlnAspValPheLeuLe 277

1783 AATAGTAGCACATTTTGAATGCAGTCCGAAGTCGCAGAGAGCCGCTGGT 1832
     ::::||||||||||||||||||||||||
907 CCAAATGGCATTGACACTCATTTTGATGAACTGCAGGATGTATTCCTAAT 956

543 rSerTrpArgHisLysGluAsnValGluGlnSerCysGluProGlyHisG 560
     |||:||    :::::::|||||         |||
1833 CTATTCGCAATTCCAGAGGCGAAATGAAGAGGCGAAAGAAGAGATCAGAG 1882
277 u......ProAspProSerGlyGlnTrpArgAspThrArgValTyrGlyTyrV 292
     |||||      |||||||||:::||
957 GAACTTTAAAGATCCT........AAAAATCCAGTTGTATATGGGAG 994

560 lnSerProAsnCysIle...............LeuPheIleGluAsnLeu 571
     |||||         ::::::||||::||:||
1883 TGGATGATCATCATCAGGACAGATCAAGGCCTTCTGCTACGTAGTCTA 1932
292 alPhe........SerAsnProTyrAsnTyrSerAlaValCysValTyr 305
     |||||         ||||||||||||||||||||:::
995 TGTTTACGACTTCCAGTAACATTTTCAAGGGATCAGCCGTGTGTATGTAT 1044

572 ThrAlaGlnGlnThrTyrGlyHisTyrPheCysGlnAlaCCAATTACCTCTGCCATGCGGTGAACATGGGTT 588
     |||::         |||||||||||:::
1933 CAACAGAAGGATTCAGGCAATTACCTCTGCCATGCGGTGAACATGGGTT 1982
306 SerLeuGlyAspIleAspValPhe............... 314
     :|||:|||     ::: :||||||
1045 AGCATGACTGATGTGAGAAGGGTGTTCCTTGGTCCATATGCCCACAGGA 1094

588 rPheArgGluAlaGlnHisTrpGlnLeuLeuProGluAspGlyIleMetA 605
     |||         |||       :::
1983 CATACAAACTCTTCTTAAGGTAACCCTG........GAAGTCATTGACA 2023
315 ....ArgThrSerLeuLysGlyTyrHisSerSerLeuProAsnProA 330
     ||||||     ::: ||||  ||||||
1095 TGGACCCAACTATCAATGGGTGCCTATCAAGGAAGAGTCCCCTATCCAC 1144

605 laGluHisLeu 608
     :||||||||
2024 CAGAGCATTTG 2034
330 rgProGlyLysCys..................LeuProAspGlnGln 339
     ||||||         ::: ||
1145 GGCCAGGAACTTGTCCCAGCAAAAACATTTGGTGGTTTTGACTCTACAAAG 1194

340 ProIleProThrGluThrPheGlnValAlaAlaAspArgHisProGluValAl 356
     ::||||   |||       ||||||  :::
1195 GACCTTCCTGATGATGTTATAACCTTTGCAAGAGTCATCCAGCCATGTA 1244

356 aGlnArgValGluProMetGly..........ProLeuLysThrProL 369
     |||||:|||:::|||     :::::|||||
1245 CAATCCAGTGTTCCTATGAACAATCGCCCAATAGTAGTCAAAACGGATG 1294

369 euPheHisSerLysTyrHisTyrGlnLysValAlaValAlaHisArgMetGln 385
     ::    :::::||||||||||:::
1295 TA........AATTATCAATTACACAATTGTCGTAGACCAGTGGAT 1335

386 AlaSerHisGlyGluTyrPheHisValLeuTyrLeuThrThrAspArgGl 402
     |||:::||   |||     ::|||||| ||
1336 GCAGAAGATGGACAG..TATGATGTTATGTTTATCGGAACAGATGTTGG 1382

402 yThrIleHisGlyGlnValGlu...ProGlyGluGlnHisGluHisSerPheA 418
     |||:::    :::::::||||  :::
1383 GACCGTTCTTAAAGTAGTTTCAATTCCTAGGGACACTTGGTATGATTTTAG 1432

418 laPheAsnIleMet...GluIleGlnProPheArgArgAlaAlaAlaIle 433
     |||:::|||  :::  |||  ||:::
1433 AAGAGGTTCTGCTGGAAGAAATGACAGTTTTCTGGGAACCGACTGCTATT 1482

434 GlnThrMetSerLeuAspAlaGluArgArgLysLeuTyrValSerSerGl 450
     :::|||:||   ::||||  ::||||||:
1483 TCACAATGGAGCTTTCCACTAAGCCACTATATATTGGTTCAAC 1532

450 nTrpGluValSerGlnValProLeuAspLeuCysGluLeuTyrGlyValTyrGlyValGlyTyrGlyG 467
     :::||||||:::::||||
```

```
seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-136-922-1

; seq_documentation_block:
; Sequence 1, Application US/08136922
Patent No. 5416197
GENERAL INFORMATION:
   APPLICANT: Raper, Jonathan A.
   APPLICANT: Luo, Yuling
   TITLE OF INVENTION: Compositions Which Regulate Neural
   TITLE OF INVENTION: Regeneration and Methods of Making and Using the Same
   NUMBER OF SEQUENCES: 2
   CORRESPONDENCE ADDRESS:
      ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
      ADDRESSEE: No. 541619rls
      STREET: One Liberty Place
      CITY: Philadelphia
      STATE: PA
      COUNTRY: USA
      ZIP: 19103
   COMPUTER READABLE FORM:
      MEDIUM TYPE: Floppy disk
      COMPUTER: IBM PC compatible
      OPERATING SYSTEM: PC-DOS/MS-DOS
      SOFTWARE: PatentIn Release #1.0, Version #1.25
   CURRENT APPLICATION DATA:
      APPLICATION NUMBER: US/08/136,922
      FILING DATE:
      CLASSIFICATION: 435
   ATTORNEY/AGENT INFORMATION:
      NAME: DeLuca, Mark
      REGISTRATION NUMBER: 33,229
```

```
;  REFERENCE/DOCKET NUMBER: UPN-1428
;  TELECOMMUNICATION INFORMATION:
;  TELEPHONE: 215-568-3100
;  TELEFAX: 215-568-3439
;  INFORMATION FOR SEQ ID NO: 1:
;  SEQUENCE CHARACTERISTICS:
;    LENGTH: 1481 base pairs
;    TYPE: nucleic acid
;    STRANDEDNESS: double
;    TOPOLOGY: both
;  MOLECULE TYPE: cDNA
;  FEATURE:
;    NAME/KEY: CDS
;    LOCATION: 50..1480
US-08-136-922-1

alignment_scores:
     Quality:  547.00        Length:     462
     Ratio:      1.913       Gaps:        17
Percent Similarity: 61.905   Percent Identity: 31.385

alignment_block:
US-09-041-236-2 x US-08-136-922-1

Align seg 1/1   to: US-08-136-922-1  from: 1  to: 1481

197 IleValHisGlnAlaTyrArgAspLysIleTyrPhePheAr       213
    ::|||||||::|:::||||||::|||:|||::||:|::::||
 47 ATACTTGAACATGACAATCCTGAAGATGACAAAGTATACTTTTCTTCCG    96

213 gGluAspAspLysAsnProAspGluAlaProLeuAsnValSerArgV   230
    |||::|   |||   :::::||::::||
 97 TGAAAATGCAATAGATGGAGAACACTCTGGAAAAGCTACTCACGCTAGAA  146

230 aAlaGlnLeuCysArgGlyAspGlnGlyGlyGluSerSerLeuSerVal  246
    :::|||||::|:||||:|||  ||||||:  |||||| ...GTG
147 TAGGTCAGATATGCAAGGACTTTGGAGGGCCACAGAAGTCTG...GTG   193

247 SerLysTrpAsnThrPheLeuLysAlaMetLeuValCysSerAspAlaAl 263
    ::|||||||:::::|||||||:::||||||||||||||
194 AATAAATGGACAACATTCCTCAAAGCTCGTCTGATTTGCTCAGTGCCAGG  243

263 aThrAsn........LysAsnPheAsnArgLeuGlnAspValPheLeuL  277
    :  |||        :::::||||||:::|||||||||||||
244 TCCAAATGGCATTGACAACTCATTTGATGAACTGCAGGATGTATTCCTAA  293

277 eu......ProAspProSerGlyGlnTrpProArgValTyrArgGly    291
    |||||        ::||||||||||||:|||||||||||||
294 TGAACTTTAAAGATCCT............AAAAATCCAGTGTATATGGA   331

292 ValPhe.........SerAsnProTrpAsnTyrSerAlaValCysValTy 305
    ||||||        ::| :|||||||||||||:::|||||||
332 GTGTTTACGACTTCCAGTAACATTTCAAGGGATCAGCCGTGTGTATGTA   381

305 rSerLeuGlyAspIleAspLysValPhe..................     314
    |||::||::::|||:::||||||
382 TAGCATGAGTGATGTGGAGAAGGGTGTTCCTTGGTCCATATGCCCACAGGG 431

315 .....ArgThrSerLeuLysGlyTyrHisSerSerLeuProAsnPro    329
    ::            :::|||| :::|||||
432 ATGGACCCAACTACCATCAATGGGCGCCCTATCAAGGAAGGAGTCCCCTATCCA 481

330 ArgProGlyLysCys.............LeuProAspGlnGl         339
    ||||||:|||||              :: ||||||||||:
482 CGGCCAGGAACTTGTCCCAGCAAAACATTTGGTGGTTTGACTCTACAAA   531

339 nProIleProThrGluThrPheGlnValAlaAspArgHisProGluValA 356
    ::||||  ::|:|::||  ||||| :::
532 GGACCTTCCTGATGATGTTATAACCTTTGCAAGAAGTCATCCAGCCATGT  581

356 laGlnArgValGluProMetGly.............ProLeuLysThrPro 368
    :::   |||||||:|:            :::||||||

582 ACAATCCAGTGTTCCTATGACAATGCCCAATGTGATCAAAACGGAT     631

369 LeuPheHisSerLysTyrHisTyrGlnLysValAlaValHisArgMetGl 385
    ::||:|:::|        :::||||
632 GTA.......AATTATCAATTACACAAATCGTCGAGACCGAGTGA       672

385 nAlaSerHisGlyGluThrPheHisValTyrLeuThrThrAspArgG    402
    :|||:||        |||::|:||||::|||
673 TGCAGAAGATGGACAG...TATGATGTTATGTTATCGGAACAGATGTTG   719

402 lyThrIleHisLysValValGlu...ProGlyGlnGluHisSerPhe    417
    |||||:|| ||||||| |||   |||  ||    ::::::::::::
720 GGACCGTTCTAAAGTAGTTTCAATTCCAAGGAGACTTGGTATGATTTA    769

418 AlaPheAsnIleMet...GluIleGlnProPheArgArgAlaAlaAlaIl 433
    ::::::||: ||        :::|||| ::|||:::
770 GAAGAGGTTCTGCTGGAAGAAATGACAGTTTTCGGGACCGACTGCTAT    819

433 eGlnThrMetSerLeuAspAlaGluArgGlyLeuTyrValSerSerG    450
    ::|:|||        ::::||:|||||||||||:|:|||||
820 TTCAGCAATGGAGCTTTCCACTAAGCAGACCCCTTACTTGTGCTTGGGATGG 869

450 lnTrpGluValSerGlnValProLeuAspLeuCysGluValTyrGlyGly 466
    ||||:|:::|        :||::::|:||||:||||||:
870 CGGCTGGGGTGCCCAGCTCCCTTACACGGTGTGATATTTACGGGAAA     919

467 GlyCysHisGlyCysLeuMetSerArgAspProTyrCysGlyTrpAspGl 483
    ::||| ||:|||||||||:||:        :::||||||
920 GGCGTGTCGAGTGTTGCCTGCCCGAGACCCTTACTTGTGCTTGGGATGG   969

483 nGlyArgCysIleSerIleTyr.....SerSerGluArgServalLeuG  498
    ||||  :||||||||        |:::|||||:||||:
970 TTCTGCAATGTTCTCGCTATTTCCCACTGCAAAGAGACGCCACAAGACGAC 1019

498 lnSerIleAsnProAlaGluProHisLysGluCysProAsnProLysPro 514
    ||||||  |||:||||||||| :::|||        ||||||
1020 AAGATATAAGAAATGGAGACCCCACTCACTGTCAGACTTACACCAT      1069

515 AspLys.........AlaProLeuGlnLysValSerLeu...AlaPr    526
    |||:::        ::::|||        :::||
1070 GATAATCACCATGGCCACAGCCCTGAAGAGAGAATCATCTATGGTGTAGA 1119

526 oAsnSerArgTyrTyrLeuSerCysProMetGluSerArgHisAlaThrT 543
    ::::::::|        |||| :::|||||
1120 GAATAGTAGCACATTTTGGATGACGAGTCCGAAGTCGCAGAGAGGCGCTGG 1169

543 yrSerTrpArgHisLysGluAsnValGlnGlnSerCysGluProGlyHis 559
    ||||||:|||:::        :::||||
1170 TCTATTGGCAATTCCAGAGGCGAAATGAAGAGCGAAAAGAAGAGATCAGA 1219

560 GlnSerProAsnCysIle..................LeuPheIleGluAsnLe 571
    :::|||||||||        ::::|||||:|
1220 GTGGATGATCATCATCAGGACAGATCAAGGCCTTCGCTACGTAGTCT    1269

571 uThrAlaGlnInTyrGlyHisTyrPheCysGluAlaGlnGluGlySerT  588
    |::::|:|:        |||:::||||
1270 ACAACAGAAGGATTCAGGCAATTACCTCTGCCATGCGGTGGAACATGGGT 1319

588 yrPheArgGluAlaGlnInHisTrpGlnLeuLeuProGluAspGlyIleMet 604
    |||:::        :::|||||
1320 TCATACAAACCTCTTCTAAGGTAACCCTGGAAGTCATTGACAACGAGCAT 1369

605 AlaGluHisLeuLeu.............GlyHis    611
    ::::::::|||||||        :::|||||
1370 TTGGAAGAACTTCTTCATAAAGTTGATGATGGACAT 1405
```

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-121-713D-57

```
        APPLICANT:   Kolodkin, Alex L.
        APPLICANT:   Matthes, David
        APPLICANT:   Bentley, David R.
        APPLICANT:   O'Connor, Timothy
        TITLE OF INVENTION:  The Semaphorin Gene Family
        NUMBER OF SEQUENCES:  100
        CORRESPONDENCE ADDRESS:
        ADDRESSEE:   SCIENCE & TECHNOLOGY LAW GROUP
        STREET:  268 Bush Street, Suite 3200
        CITY:  San Francisco
        STATE:  CA
        COUNTRY:  USA
        ZIP:  94104
        COMPUTER READABLE FORM:
        MEDIUM TYPE:  Floppy disk
        COMPUTER:  IBM PC compatible
        OPERATING SYSTEM:  PC-DOS/MS-DOS
        SOFTWARE:  PatentIn Release #1.0, Version #1.25
        CURRENT APPLICATION DATA:
        APPLICATION NUMBER:  US/08/121,713D
        FILING DATE:  13-SEP-1993
        CLASSIFICATION:  514
        ATTORNEY/AGENT INFORMATION:
        NAME:  Osman, Richard A.
        REGISTRATION NUMBER:  36,627
        REFERENCE/DOCKET NUMBER:  B94-002-1
        TELECOMMUNICATION INFORMATION:
        TELEPHONE:  (415)343-4341
        TELEFAX:  (415) 343-4342
        TELEX:
        INFORMATION FOR SEQ ID NO:  57:
        SEQUENCE CHARACTERISTICS:
        LENGTH:  2854 base pairs
        TYPE:  nucleic acid
        STRANDEDNESS:  double
        TOPOLOGY:  linear
        MOLECULE TYPE:  cDNA
        FEATURE:
        NAME/KEY:  CDS
        LOCATION:  451..2640
US-08-121-713D-57

alignment_scores:
        Quality:     449.00              Length:       569
        Ratio:        1.448              Gaps:          28
Percent Similarity:  54.482    Percent Identity:  27.592

alignment_block:
US-09-041-236-2 x US-08-121-713D-57

Align seq 1/1  to:  US-08-121-713D-57  from: 1  to: 2854
```

```
    4 LeuLeuTrpAlaAlaAlaAlaSerAlaGlnGlyHisLeuArgSerGlyPr  20
      |||||||||||||||::|||:::|||:||
  481 CTGCTTTGGGTGGCCGCTGCACGCCCGCCGCATGGGTCAACGACGTCAGCCC 530

   20 oArgIlePheAlaValTrpLysGlyHisValGlyGlnAspArgValAsp.  36
      |:|||:::|||||||:::||::|| |||::::
  531 CAAGATGACGTC........CAGTTCGGTGAGGAACGCGGTGCAAC 568

   37 ..PheGlyGlnThrGluProHisThrValLeuPheHis.....GluPro  50
      ||| :|||:|||  |||    |||
  569 GCTTCCTGGCAATGAATGACACAAAGACCACCTTCAAGCTGCTGGAGAAG 618

   51 GlySerSerValTrpValGlyArgGlyLysValTyr.........  64
      ||||||:: |||||||::|||:::||::
  619 GACCACAACTCGCTCCTCGTAGGAGCTAGGAACATCGTCTACAATATCAG 668

   65 .LeuPheAspPheProGluGlyLysAsnAlaSerValArgThrValAsnI  81
      ||| ||| |||||| :|||||    :::
  669 CCTTCGAGACCTCACAGAATTCACCGAGCAGAGGATCGAGTGGCACTCGT 718
```

```
 81 leGlySerThrLysGlySerCysLeu...........AspLysArgAsp  93
    ||||||                       :::||| |||||
719 CAGGTGCCCATCGCGAGCTCTGCTACCTCAAGGGGAAGTCAGAGGACGAC 768

 94 CysGluAsnTyrIleThrLeuLeuGluArg...ArgSerGluGlyLeuLe 109
    |||:|||||||||:||  :::|||  ::||||   :::|| :|||
769 TGCCAGAACTACATCCGAGTCCTGGGGAAAATTGACGATGACCGCGTACT 818

109 uAlaCysGlyThrAsnAlaArgHisProSerCys.....TrpAsnLeuV 124
    |||||||||||:|||||:: ||| |||    |||
819 CATCTGCGGTACGAACGCCTATAAGCCACTATGTCGGCACTACGCCCTCA 868

124 alAsnGlyThrValProLeuGlyGluMet......ArgGlyTyrAla 138
    :::|||                        |||||| |||
869 AGGATGGAGATTATGTTGTAGAGAAGAAATATGAGGGAAGAGGATTGTGC 918

139 ProPheSerProAspGluAsnSerLeuValLeuPheGluGluAspGluVa 155
    ||||:|||::|||||:||||:|      :::::
919 CCATTTGACCCTGACCAACAGCACTGCAATATACAGTGAGGGACAATT 968

155 lTyrSerThrIleArgLysGlyGlnGluTyrAsnGlyIleLysIleProArgPheA 172
    :||| |||:||                        |||||:|||
969 GTACTCAGCA...ACAGTGGCAGACTCTCTGGAACTGACCCTCTCATA 1014

172 rgArgIleArgGlyGluSerGluLeuTyrThrSerAspThrValMetGln 188
    |||||||                        ||||||
1015 .....TACCCGCGGCCCTCTAAGACAAGCAGAGATCTGACCTCAAACAATTA 1059

189 Asn...ProGlnPheIleLysAlaThrIleValHisGlnAsnGlnAlaTy 204
    |||     |||:::|||||||||                   ||
1060 AATGCTCCTAACTTTGTCAACACA...........ATGGAGTA 1091

204 rAspAspPheIleTyrTyrPheArgGluAspAsnProAspLysAsnP 221
    :|||:||:|||:||||||:::|||||
1092 CAATGATTTTATATTTTCTTCCGGAGACTGCGTGTTGAGTACATCA 1141

221 roGluAlaProLeuAsnValSerArgValAlaGlnLeuCysSerArgGlyAsp 237
    |||||||||||||||||::|||||:::|||||||
1142 ACTGCGGAAAGGCTATCTATTCAAGAGTTGCCAGAGTCTGTAAAACATGAC 1191

238 GlnGlyGlyGluSerSerLeuSerValSerLysTrpAsnThrPheLeuLy 254
    :||||||                        ::||||||||||||
1192 AAGGGGCGGCCCTCATCAG...GGTGGTGACAGATGGACTTCTTTTTGAA 1238

254 sAlaMetLeuValCysSerAspAlaAlaThrAsnLys...AsnPheAsnA 270
    |:::|||                        ||||||
1239 ATCACGTCTGAACTGTTCCGTCCCTGGAGATTATCCATTTTACTTCAATG 1288

270 rgLeuGlnAspValPhe...LeuLeuProAspProSerGlyGlnTrpArg 285
    |:::||:|||::|||   |:::|||
1289 AAATTCAGTCAACAAGTGACATCATTGAAGGAAATTATGGTGGTCAAGTG 1338

286 AspThrArgValTyrGlyValPheSerAsnProTrpAsnTyr......... 299
    :::|||||||||||:::::||| |||
1339 GAGAAACTCATCTACGGTGTCTTCAGCAACCCAGTGAACTCTATTGGTGG 1388

300 .SerAlaValCysValTyrSerLeuGlyAspIleAspLysValPhe..... 314
    |||||||:::|||  :::|||
1389 CTCTGCTGTTTGTGCCTTCAGTATGAAGTCAATACTTGAGTCATTTGATG 1438

315 ..........................ArgThrSer 317
1439 GTCCATTTAAAGAGGAGGAAAACGATGAACTCAAACTGGTTGGCAGTGCCA 1488

318 SerLeuGlyTyrHisSerSerLeuProAsnProArgProGlyLysLysCy 334
    ||||||||                        :::||||||||:||:
1489 AGCCCTAAA...........GTGCCAGAACCAAGGCCTGGACAATG 1523

334 sLeuProAspGlnGlnProIleProThrGluThrPheGlnValAlaAspA 351
    :::| |||::::::   :::|||
1524 TGTGAATGACAGTCGTACACTTCCTGATGTGTCTGCAATTTGTAAAGT 1573

351 rgHisProGluValAlaGlnArgValGluProMetGlyProLeuLysThr 367
```

APPLICATION NUMBER: US/08/835,268
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/121,713
FILING DATE: 13-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A.
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: B94-002-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)343-4341
TELEFAX: (415) 343-4342
TELEX:
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 2854 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 451..2640
US-08-835-268-57

alignment_scores:
Quality: 449.00   Length: 569
Ratio: 1.448   Gaps: 28
Percent Similarity: 54.482   Percent Identity: 27.592

alignment_block:
US-09-041-236-2 x US-08-835-268-57

Align seg 1/1 to: US-08-835-268-57 from: 1 to: 2854

seq_documentation_block:
; Sequence 57, Application US/08835268
; Patent No. 5807826
GENERAL INFORMATION:
APPLICANT: Goodman, Corey S.
APPLICANT: Kolodkin, Alex L.
APPLICANT: Matthes, David
APPLICANT: Bentley, David R.
APPLICANT: O'Connor, Timothy
TITLE OF INVENTION: The Semaphorin Gene Family
NUMBER OF SEQUENCES: 100
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 Bush Street, Suite 3200
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

seq_name: /cgn2_6/ptodata/2/1na/5B_COMB.seq:US-08-835-268-57

```
139 ProPheSerProAspGluAsnSerLeuValLeuPheGluGlyAspGluVa 155
    |||||:::||::||||||||              :::||||
919 CCATTTGACCCTGACCAACAGCACTGCATATACAGTGAGGGACAATT 968

155 lTyrSerThrIleArgLysGlnGluTyrAsnGlyLysIleProArgPheA 172
    :||||||:::        ::|||||||||||   ||| :::
969 GTACTCAGCA...ACAGTGGCAGACTTCTCTGGACTGACCCTCTCATA. 1014

172 rgArgIleArgGlyGluSerGluLeuTyrThrSerAspThrValMetGln 188
    |||||        ||||||              |||||: ||| :::
1015 ....TACCCGGGCCCTCTAAGAACAGAGAGATCTGACCTCAAACAATTA 1059

189 Asn...ProGlnPheIleLysAlaThrIleValHisGlnAsnGlnAlaTy 204
    |||  ||||::::||:::::::||            ::|||
1060 AATGCTCCTAACTTTGTCAACACA...........ATGGAGTA 1091

204 rAspAspLysIleTyrTyrPheArgPheArgAspAsnProAspLysAsnP 221
    |:||||   |||:::|||||||||||              :::
1092 CAATGATTTTATAATTCTTCTTCTTCCGAGAGACTGCTGTTGAGTACATCA 1141

221 roGluAlaProLeuAsnValSerArgValAlaAlaGlnLeuCysArgGlyAsp 237
    :::     |||||||||||||||||              |||
1142 ACTGCGGAAAGGCTATCTATTCAAGAGTTGCCAGAGTCTGTAAACATGAC 1191

238 GlnGlyGlyGluSerSerLeuSerValSerLysTrpAsnThrPheLeuLy 254
    ::||||||  |||||           ::|||||||||||||||
1192 AAGGGCGGGCCCTCATCAG...GGTGGTGACAGATGGACTTCTTTTTTGAA 1238

254 sAlaMetLeuValCysSerSerAspAlaAlaThrAsnLys...AsnPheAsnA 270
    |:::: ||||||              |||||||
1239 ATCACGTCTGAACTGTTCCGTCCGGAGATTATCCATTTACTTCAATG 1288

270 rgLeuGlnAspValPhe...LeuLeuProAspProSerGlyGlnTrpArg 285
    :|||:||:||              ||
1289 AAATTCAGTCAACAGTGACATCATTGAAGGAAATTATGTGGTCAAGTG 1338

286 AspThrArgValTyrGlyValPheSerAsnProTrpAsnTyr........ 299
    :::        ::|||||||||||
1339 GAGAAACTCATCTACGGTGTCTTCACGACCACCAGTGAACTCTATTGGTGG 1388

300 .SerAlaValCysValTyrSerLeuGlyAspIleAspLysValPhe..... 314
    |||||||||||:::::::::||   :::  |||
1389 CTCTCGCTGTTTGTGCCTCAGTATGAAGTCAATACTTGAGTCATTTGATG 1438

315 ...................................ArgThrSer 317
    :::
1439 GTCCATTTAAAGAGCAGGAGAAACGATGAACTCAAACTGGTTGGCAGTGCCA 1488

318 SerLeuLysGlyTyrHisSerSerLeuProAsnProAsnProGlyLysCy 334
    |||||||              ::|||||||||||:::|
1489 AGCCTTAAA................GTGCCAGAACCAAGGCCTGGACAATG 1523

334 sLeuProAspGlnGlnProThrGluThrPheGlnValAlaAlaAspA 351
    |:::        ||:::||||    ::: :::
1524 TGTGAATGACAGTCGTACACTTCCTGATGTGTCGTCAATTTGTAAAGT 1573

351 rgHisProGluValAlaGlnArgValGluProMetGlyProLeuLysThr 367
    |||||||||||              :::
1574 CACATACACTGATGATGAGGCCGTG.......................... 1599

368 ProLeuPheHisSerLys.........................TyrHi 375
    |||  |||              |||:|
1600 CCAGCATTTTTACTCGGCCAATTCCATTCGGATCAGTTACAGTACAG 1649

375 sTyrGlnLysValAlaVal...HisArgMetGlnAlaSerHisGlyGluT 391
    :::: |||:||||              |||:||||
1650 ATTTACAAAAATAGCTGTTGATCAACAAGTCCGAACACCAGATGGGAAAG 1699

391 hrPheHisValLeuThrThrAspArgGlyThrIleHisLysVal 407
    :||||| :::              |||:::|
1700 CGTATGATGTCCTGTTTATAGGAACTGATGATGGCAAAGTGATAAAAGCT 1749

408 ValGluProGlyGluGlnGluHisSerPhePheAlaPheAsn........Il 421
    :::
```

(right column)

```
    :::::    :::::              |||
1750 TTGAACTCTGCCTCCTTGATTCATCTGATACTGATAGTGTGTAAT 1799

421 eMetGluIleGlnProPheArgArgAlaAlaAlaIleGlnThrMet.... 436
    |  :|||::||:::              |||
1800 AGAGAACTGCCAAGTGTGCCACCTGGAGTACCTGTTAAGAACCTGTATG 1849

437 .....SerLeuAspAlaGluArgArgLysLeuTyrValSerSerGlnTrp 451
    ::::||||||:||:              ::|||  |||
1850 TGGTGCGAATGGATGGGGATGATAGCAAGCTGGTGGTTGTGTCGATGAT 1899

452 GluValSerGlnIleValProLeuAspLeuCysValTyrGlyValTyrGly...... 466
    |||:|||:::              ::|||
1900 GAGATCTGGCAATTAAGCTTCATCGTGT.........GGCTCAGATAA 1940

467 ........GlyCysHisGlyCysLeuMetSerArgAspProTyrCysGlyT 481
    ::::::||||||||||:              ::::||||||||||
1941 AATAACAAATTGTCAGAAGTGTGTCCTTGCAAGATCCTTACTGTGCAT 1990

481 yrAsp.......GlnGlyArgCysIleSerIleTyrSer.......... 491
    ||||      :::|||              |||
1991 GGGACAATGTAGAATTAAAATGTACAGCTGTAGGTTCACCAGACTGGAGT 2040

492 ...SerGluArgSerValLeuGlnSerIleAsnProAlaGluProHisLy 507
    ::|||||||||              :::|| |||
2041 GCTGGAAAAGCGCTTATTCAGAACATTTCACTCGGTGAA...CATAA 2087

507 sGluCys 509
    | |||
2088 AGCTTGT 2094
```

seq_name: /cgn2_6/ptodata/2/ina/5D_COMB.seq:US-09-060-692-57

```
;   LENGTH: 2854 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
;   FEATURE:
;     NAME/KEY: CDS
;     LOCATION: 451..2640
US-09-060-692-57

alignment_scores:
          Quality:  449.00          Length:    569
          Ratio:      1.448            Gaps:       28
Percent Similarity: 54.482    Percent Identity: 27.592

alignment_block:
US-09-041-236-2 x US-09-060-692-57

Align seg 1/1  to: US-09-060-692-57  from: 1  to: 2854
```

```
    4  LeuLeuTrpAlaAlaAlaAlaSerAlaGlnGlyHisGlyLeuArgSerGlyPr  20
       ||||||:::||:::|||:::|||         :::||
  481  CTGCTTTGGGTGGCGCTGCGCCGCCGCACGCGCCGATGGGTCAACGACGTCAGCCC  530

   20  oArgIlePheAlaValTrpLysGlyHisValGlyGlnAspArgValAsp.  36
       |:::::::::::||         :::|||:::::|||:::||:::
  531  CAAGATGTACGTC..........CAGTTCGGTGAGGAACGGGTGCAAC  568

   37  ..PheGlyGlnThrGluProHisThrValLeuPheHis......GluPro  50
         |||  :::|||  |||         |||
  569  GCTTCCTGGCCAATGAATCGCCACAAAGACCACTTCAAGCTGCTGGAGAAG  618

   51  GlySerSerValTrpValGlyGlyArgGlyLysValTyr.........  64
       :::::||  ||||||:::|||||:::|||:::|||         |
  619  GACCACAACTCGCTCCTCGTAGGAGCTAGGAACATCGTCTACAATATCAG  668

   65  .LeuPheAspPheProGluGlyLysAsnAlaSerValArgThrValAsnI  81
       |||  ||||:::||         :::|||
  669  CCTTCGAGACCTCACAGAATTCACCGAGCAGAGGATCGAGTGGCACTCGT  718

   81  leGlySerThrLysGlySerCysLeu..........AspLysArgAsp  93
       :::|||:::  :::|||         :::|||  |||
  719  CAGGTGCCCATCGCGAGCTCTGCTACCTCAAGGGGAGTCAGGGACGAC  768

   94  CysGluAsnTyrIleThrLeuLeuGluArg...ArgSerGluGlyLeuLeu  109
       |||||||||:::|||  ||:::  :::|||   :::|||
  769  TGCCAGAACTACATCCGAGTCCTGGCGAAATTGACGATGACCGCGTACT  818

  109  uAlaCysGlyThrAsnAlaArgHisProSerCys......TrpAsnLeuV  124
       :::||||:::|||:::||||||||||||         :::|||
  819  CATCGTGCGGTACGGAACGCCTATAAGCCACTATGTCGGCACTACGCCCTCA  868

  124  alAsnGlyThrValProLeuGlyGluMet......ArgGlyTyrAla  138
       :::||||||  :::|||||||         :::|||  :::
  869  AGGATGGAGATTATGTTGTAGAGAAGAATATGAGGGAAGAGGGATTGTGC  918

  139  ProPheSerProAspGluAsnSerLeuValLeuPheGluGlyValAspGluVa  155
       :::||||||||||||||:::|||:::|||         |||||
  919  CCATTTGACCCTGACAACAACAGCACTGCAATATACAGTGAGGGACAATT  968

  155  lTyrSerThrIleArgLysGlnGluTyrAsnGlyLysIleProArgArgPheA  172
       :::|||||:::         :::|||:::|||
  969  GTACTCAGCA...ACAGTGGCAGACTTCTCTGGAACTGACCCTCTCATA.  1014

  172  rgArgIleArgGlyGluSerGluTyrThrSerAspThrValMetGln  188
       :::|||  :::|||:::||||:::||:::|||:::
 1015  ....TACCGCCGCCCTCTAGAACACAGAGATCTGACCTCAAACAATTA  1059

  189  Asn...ProGlnPheIleLysAlaThrIleValHisGlnAspGlnAlaTy  204
        |||   :::||||:::::::::||:::         :::  |||
 1060  AATGCTCCTAACTTTGTCAACACA........................ATGGAGTA  1091
```

```
  204  rAspAspLysIleTyrTyrPheArgGluAspAsnProAspLysAsnP  221
       |:::|||||| |||:::|||||||         :::
 1092  CAATGATTTATATTCTTCTTCCGAGAGACTGCTGTTGAGTACATCA  1141

  221  roGluAlaProLeuAsnValSerArgValAlaGlnLeuCysArgGlyAsp  237
       :::||||||:::|||||||:::|||:::::::::|||||||:::
 1142  ACTGCGGAAAGGCTATCTATTCAAGAGTGCCAGAGTCGTAAACATGAC  1191

  238  GlnGlyGluSerSerValSerLysTrpAsnThrPheLeuU  254
       :::|||||         :::|||||||||||||||:::
 1192  AAGGGCGGCCCTCATCAG.....GGTGGTGACGATGGACTTCTTTTTGAA  1238

  254  sAlaMetLeuValCysSerAspAlaAlaThrAsnLys...AsnPheAsnA  270
       |:::  :::|||  :::  ::.||         ||||||  ||::::
 1239  ATCACGTCTGAACTGTTCCGTCCCTGGAGATTATCCATTTACTTCAATG  1288

  270  rgLeuGlnAspValPhe...LeuLeuProAspProSerGlyGlnTrpArg  285
       :::|||:::|||::         :::||||||||||||:::|||||
 1289  AAATTCAGTCAACAAGTGACATCATTGAAGGAAATTATGGTGGTCAAGTG  1338

  286  AspThrArgValTyrGlyValPheSerAsnProTrpAsnTyr.......  299
       :::  :::||||||:::|||||||||||||||::::::
 1339  GAGAAACCATCATCACGGTGTCTTCACGACACCCAGTGAACTCTATTGGTGG  1388

  300  .SerAlaValCysValTyrSerLeuGlyValAspIleAspLysValPhe....  314
       |||||||:::|||||||         :::|||
 1389  CTCTGCTGTTGTGCCTTCAGTATGAGTCAATACTTGAGTCATTTGATG  1438

  315  ............................................ArgThrSer  317
                                                     :::
 1439  GTCCATTTAAAGAGCAGGAAACGATGAACTCAAACTGGTTGGCAGTGCCA  1488

  318  SerLeuLysGlyTyrHisSerSerLeuProAsnProArgProGlyLysCy  334
       |||||||||         :::||||||||||||||:::||:::|
 1489  AGCCTTAAA.............GTGCCAGAACCAAGGCCTGGACAATG  1523

  334  sLeuProAspGlnGlnProIleProThrGluThrPheGlnValAlaAspA  351
       :::||||||||||||||:::|||         :::|||
 1524  TGTGAATGACAGTCGTACACTTCCTGATGTGTCGTCAATTTGTAAAGT  1573

  351  rgHisProGluValAlaGlnArgValGluProMetGlyProLeuLysThr  367
       |||  ||:::
 1574  CACATACACTGATGGATGAGGCCGTG..........................  1599

  368  ProLeuPheHisSerLys..............................TyrHi  375
       ||||||  :::||         |||:::
 1600  CCAGCCATTTTTACTCGGCCAATTCTCATTCGGATCAGCTTACAGTACAG  1649

  375  sTyrGlnLysValAlaVal...HisArgMetGlnAlaSerHisGlyGluT  391
       :::|||         |||:::||||||:::|||
 1650  ATTACAAAAATAGCTGTTGATCAACAAGTCCGAACACCAGATGGGAAAG  1699

  391  hrPheHisValLeuTyrLeuThrThrAspArgGlyThrIleHisLysVal  407
       :::|||         |||:::|||
 1700  CGTATGATGTCCGTGTTTATAGGAACTGATGATGGCAAAGTGATAAAAGCT  1749

  408  ValGluProGlyGluGlnHisSerPheAlaPheAsn.........I  421
       :::|||         |||
 1750  TTGAACCTCGCCACCTTTGATTCATCTGATACTGTAGATAGTGTGTAAT  1799

  421  eMetGluIleGlnProPheArgArgAlaAlaAlaIleGlnThrMet....  436
       :::||||:::|||||         :::
 1800  AGAAGAACTGCAAGTGTTGCCACCTGGAGTACCTGTTAAGAACCTGTATG  1849

  437  .....SerLeuAspAlaGluArgLysLeuTyrValSerSerGlnTrp  451
       |||||         :::|||:::|||
 1850  TGGTTGCGAATGGATGGGGATGAGATAGCAAGCTGGTGGTGTCTGATGAT  1899

  452  GluValSerGlnIleValProLeuAspLeuCysGlnValTyrGlyGly......  466
       |||  :::||         ||
 1900  GAGATTCTGGCAATTAAGCTTCATCGTTGT.........GGCTCAGATAA  1940

  467  ..........GlyCysHisGlyCysLeuMetSerArgAspProTyrCysGlyT  481
 ```

```
  20 oArgIlePheAlaValTrpLysGlyHisValGlyGlnAspArgValAsp.          36
     |:::::::|:::      |||:::          |||:||:|||||:::|
 531 CAAGATGTACGTC............CAGTTCGGTGAGGAACGGGTGCAAC          568

  37 ..PheGlyGlnThrGluProHisThrValLeuPheHis......GluPro          50
          :::||:|:||||||:|::||          |||
 569 GCTTCCTGGGCAATGATCGCACAAAGACCACTTCAAGCTGCTGGAGAAG          618

  51 GlySerSerValTrpValGlyGlyArgGlyLysValTyr..........          64
     |:||||:|  |:||||   |::|||    |||||
 619 GACCACAACTCGCTCCTCGTAGGAGCTAGGAACATCGTCTACAATATCAG          668

  65 .LeuPheAspPheProGluGlyLyLysAsnAlaSerValArgThrValAsnI          81
      |||  |||||:::          ::||:||||::::::::
 669 CCTTCGACACCTCACAGAATTCACCGAGCAGAGGATCGAGTGGCACTCGT          718

  81 leGlySerThrLysGlySerCysLeu..........AspLysArgAsp          93
     |:|||  ::|  |||||  :::          :||||||  :::
 719 CAGGTGCCCATGCGAGCTCTGTACCTCAAGGGAGAGTCAGAGGACGAC          768

  94 CysGluAsnTyrIleThrLeuLeuGluArg...ArgSerGluGlyLeuLe          109
     |||::|:|||:|:||||          ||||
 769 TGCCAGAACTACATCCGAGTCCTGGCGAAAATTGACGATGACCGCGTACT          818

 109 uAlaCysGlyThrAsnAlaArgHisProSerCys......TrpAsnLeuV          124
     |:|||||||||||          :||  |||          :||  |||
 819 CATCTGCGGTACGAACGCCTATAAGCCACTATGTCCGCACTACGCCCTCA          868

 124 alAsnGlyThrValProLeuGlyMet........ArgGlyTyrAla          138
     ::|||          |||          |:||||  ::::||
 869 AGGATGGAGATTATGTTGTAGAGAAGAAATATGAGGGAAGAGGATTGTGC          918

 139 ProPheSerProAspGluAsnSerLeuValLeuPheGluGlyValAspGluVa          155
     |||||:||||:||::|||          |||||||||::::
 919 CCATTGACCCTGACCACAACAGCACTGCAATATACAGTGAGGGACAATT          968

 155 lTyrSerThrIleArgGlyGlnGluTyrAsnGlyLysIleProArgPheA          172
     :|:||:|::          :|:|:||::|||:||  :||
 969 GTACTCAGCA...ACAGTGGCAGACTTCTCTGGAACTGACCCTCTCATA.          1014

 172 rgArgIleArgGlySerGluLeuTyrThrSerAspThrValMetGln          188
      |:||||          |||          ::|||
1015 ......TACCGCGCCCCTCAAGAACAGAGAGATCTGACCTCAAACAATTA          1059

 189 Asn...ProGlnPheIleLysAlaThrIleValHisGlnAspGlnAlaTy          204
     |||          :|:||:|||||
1060 AATGCCTCCTAACTTTGTCAACACA...........ATGGAGTA          1091

 204 rAspAspLysIleTyrPhePheArgGluAspAsnProAspLysAsnP          221
     |:::||  |||:||  ::||||||||||
1092 CAATGATTTATATTCTTCTTCCGAGAGACTGCTGTTGAGTACATCA          1141

 221 roGluAlaProLeuAsnValSerArgValAlaLaGlnLeuCysArgGlyAsp          237
     ::  ||||||||||::|:||:|:||||:||::|:|:|||
1142 ACTGCGGAAAGGCTATCTATTCAAGAGTTGCCAGAGTCTGTAAACATGAC          1191

 238 GlnGlyGlyGluSerSerLeuSerValSerLysTrpAsnThrPheLeuLy          254
     ::||:||          ||||  |||||||||||
1192 AAGGGCGGCGCCCTCATCAG...GGTGGTGACAGATGGACTTCTTTTTGAA          1238

 254 sAlaMetLeuValCysSerAspAlaAlaThrAsnLys...AsnPheAsnA          270
     |:::  ||  ||:||||          ||||||||||||
1239 ATCACCGTCGAACTGTTCCGTCCCTGGAGATTATCCATTTTACTTCAATG          1288

 270 rgLeuGlnAspValPhe...LeuLeuProAspProSerGlyGlnTrpArg          285
     :|:::||||          :|:|
1289 AAATTCAGTCAACAGTGACATCATTGAAGGA\ATTATGGTGGTCAAGTG          1338

 286 AspThr=ArgValTyrGlyValPheSerAspAsnProTrpAsnTyr.........          299
     |:::||||          :::||
1339 GAGAAACTCATCTACGGTGTCTTCACGACACCAGTGAACTCTATTGTGG          1388
```

```
1941 AATAACAAATTGTCGAGAATGTGTGTCCTTGCAAGATCCTTACTGTGCAT          1990
     :::||||||||  |||:::          ::|||||||||||||:::|
 481 rpAsp.....GlnGlyArgCysIleSerIleTyrSer.............          491
      ||||          |::          |||
1991 GGGACAATGTAGAATTAAAATGTACAGCTGAGGTTCACCAGACTGGAGT          2040

 492 ...SerGluArgSerValLeuGlnSerIleAsnProAlaGluProHisLy          507
     ::||||||          ::|||||:|||||  |||:||          ||||
2041 GCTGGAAAAGAGACGCCTTATTCAGAACATTCACTCGGTGAA...CATAA          2087

 507 sGluCys 509
2088 AGCTGT 2094
```

seq_name: /cgn2_6/ptodata/2/1na/PCTUS9_COMB.seq:PCT-US94-10151A-57

alignment_scores:
     Quality: 449.00   Length: 569
     Ratio: 1.448   Gaps: 28
Percent Similarity: 54.482   Percent Identity: 27.592

alignment_block:
US-09-041-236-2 x PCT-US94-10151A-57

Align seq 1/1 to: PCT-US94-10151A-57 from: 1   to: 2854

```
   4 LeuLeuTrpPleAlaAlaAlaSerAlaGlnGlyHisLeuArgSerGlyPr          20
     |||||||||:|:::|||          :|:||
 481 CTGCTTTGGGTGGGCCTGCACGCCCCCGCATGGGTCAAGCGACGTCAGCCC          530
```

```
;     APPLICANT:  Bentley, David R.
;     APPLICANT:  O'Connor, Timothy
;     TITLE OF INVENTION:  The Semaphorin Gene Family
;     NUMBER OF SEQUENCES:  100
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  SCIENCE & TECHNOLOGY LAW GROUP
;       STREET:  268 Bush Street, Suite 3200
;       CITY:  San Francisco
;       STATE:  CA
;       COUNTRY:  USA
;       ZIP:  94104
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/121,713D
;       FILING DATE:  13-SEP-1993
;       CLASSIFICATION:  514
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Osman, Richard A.
;       REGISTRATION NUMBER:  36,627
;       REFERENCE/DOCKET NUMBER:  B94-002-1
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (415)343-4341
;       TELEFAX:  (415) 343-4342
;       TELEX:
;     INFORMATION FOR SEQ ID NO:  63:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  2504 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  355..2493
US-08-121-713D-63
```

```
alignment_scores:
          Quality:       408.00          Length:        517
          Ratio:           1.365         Gaps:           23
Percent Similarity:  57.834      Percent Identity:  26.692

alignment_block:
US-09-041-236-2 x US-08-121-713D-63

Align seg 1/1  to: US-08-121-713D-63    from: 1  to: 2504

 43 HisThrValLeuPheHisGluProGlySerSerValTrpValGlyGl  59
    :||  |||  :|||||||||||||||  ||||||||||
496 CACTTTATTGTCTTGAATCAA...GACGAAACTTCGATATTAGTAGGCGG  542

 59 yArgGlyLysValTyr......LeuPheAspPheProGluGlyLysA   73
    |||:||:::||||||      :|||||||||||  |||||||||
543 TAGAAATAGGGTTTACAATTAAGTATATTCGACCCTCAGTGAGCGTAAAG  592

 73 snAlaSerValArgThrValAsnIleGlySerThrLysGlySerCysLeu  89
    :|||  :||  ::|  |||  :|  ::|
593 GGGGGCCGAATCGACTGGCCATCGTCCGATGGCACATGGCCAGTTGTGTATA  642

 90 AspLys.........ArgAspCysGluAsnTyrIleThrLeuLeuGl  102
    |||          ||||||||:|||||||||:||||
643 TTGAAAGGGAAAAACGGACGACGACTGCCAAAATTACATTAGAATACTGTA  692

102 uArgArgSerGluGly..LeuLeuAlaCysGlyThrAsnAlaArgHisP  118
    |||  |:|  :||||||  ||||||||||||::|||
693 CTCTTCAGAACCGGGGAAATTAGTTATTTGCGGGACCAATTCGTACAAAC  742

118 roSerCys......TrpAsnLeuValAsnGlyThrValValProLeuGly  132
    ||  |||       ::|||  |||
```

```
 300 .SerAlaValCysValTyrSerLeuGlyAspIleAspLysLysValPhe....  314
     |||||||||||||||:::|||  :: |||  :
1389 CTCTGCTGTTTGTGTTTTGCCTTCAGTATGATGAAGTCAATACTTGAGTCATTTGATG  1438

 315 ...............................................ArgThrSer  317
                                                     :::|
1439 GTCCATTAAAGAGCAGGAAACGATGAACTCAAACTGGTTGGCAGTGCCA  1488

 318 SerLeuLysGlyTyrHisSerSerLeuProAsnProArgProGlyLysCy  334
     |||||||||            ::||||::||||||||||||||||:::||
1489 AGCCTTAAA........GTGCCAGAACCAAGGCCTGGACAATG  1523

 334 sLeuProAspGlnGlnProIleProThrGluThrPheGlnValAlaAspA  351
     |::: ::|||||            |||::::  |||:|||:|||:||||||
1524 TGTGAATGACAGTCGTACACTTCCTGATGTGTCTGTCAATTTGTAAAGT  1573

 351 rgHisProGluValAlaGlnValGluProMetGlyProLeuLysThr  367
     |||  :::  |||                            |||
1574 CACATACACTGATGGATGAGGCCGTG.......................  1599

 368 ProLeuPheHisSerLys..........................TyrHi  375
     |||  |||  ::::                              |||  |
1600 CCAGCATTTTTTACTCGGCCAATTCTCATTCGGATCAGCTTACAGTACAG  1649

 375 sTyrGlnLysValAlaVal...HisArgMetGlnAlaSerHisGlyGluT  391
     |:::  ||||||||||||     |||::|||  ||  |||||||  |||
1650 ATTTACAAAAATAGCTGTTGATCAACAAGTCCGAACACCAGATGGGAAAG  1699

 391 hrPheHisValLeuTyrLeuThrThrAspArgGlyThrIleHisLysVal  407
     |:::|||  ||||||::|||::||:||||||  :::  ||||
1700 CGTATGATGTCCTGTTTATAGGAACTGATGATGGCAAAGTGATAAAAGCT  1749

 408 ValGluProGlyGluGlnGluHisSerProPheAlaPheAsn........Il  421
     |:::  ::|||  :::  :::||||  :::  :::|||||          |
1750 TTGAACCTGCCTCCTTGATCATCTGATACTGTAGATAGTGTTGTAAT  1799

 421 eMetGlyIleGlnProPheArgArgAlaAlaAlaAlaIleGlnThrMet....  436
     |:::||||||||||  :::||  :|||||||||||||||||:::::|
1800 AGAAGAACTGCAAGTGTTGCCACCTGGAGTACCTGTTAAGAACCTGTATG  1849

 437 .....SerLeuAspAlaGluArgArgLysLeuTyrValSerSerGlnTrp  451
          ::|||||||:::|||||||||||||||||||||||||  |||:::|
1850 TGGTGCCAATGGAATGGGGATGAGTAGCCAAGCTGGTGGTTGTGTCTGATGAT  1899

 452 GluValSerGlnValProLeuAspLeuCysGluValTyrGlyGly......  466
     ||||::||:||    :::|||  ||  |||  :::||||||||
1900 GAGATTTCGGCAATTAAGCTTCATCGTGT........GGCTCAGATAA  1940

 467 .......GlyCysHisGlyCysLeuMetSerArgAspProTyrCysGlyT  481
            ||||::||  ::|||||  :::|||||  :|||||  |||||||
1941 AATAACAAATTGTCGAGAATGTGTCGTCCTTGCAGAGTCCTTACTGTGCAT  1990

 481 rpAsp......GlnGlyArgCysIleSerIleTyrSer.............  491
     |::||       ||||  |||||||  :::  |:::::|
1991 GGGACAATGGAATTAAAATGTAGACAGCTGTAGGTTCACCAGACTGGAGT  2040

 492 ...SerGluArgSerValLeuGlnSerIleAsnProAlaGluProHisLy  507
        ::::||  |||::: |||        :::||  ||:::||  |||||
2041 GCTGGAAAAAGACACGCTTTATTCAGAACATTTCACTCGGTGAA...CATAA  2087

 507 sGluCys  509
     ||  |||
2088 AGCTTGT  2094
```

```
seq_name: /cgn2_6/ptcdata/2/lna/5A_COMB.seq:US-08-121-713D-63

seq_documentation_block:
; Sequence 63, Application US/08121713D
; Patent No. 5639856
;     GENERAL INFORMATION:
;     APPLICANT:  Goodman, Corey S.
;     APPLICANT:  Kolodkin, Alex L.
;     APPLICANT:  Matthes, David
```

```
743   CCCTCTGTCCGACGTACGCATTAAGGAGGAAAGTACCTGGTTGAGAAA  792

133   GluMetArg......GlyTyrAlaProPheSerProAspGluAsnSerLe  147
        ||||||    |||  ||    || ::   |||||||:||||||
793   GAAGTAGAAGGGATAGGCTTGTCCCATACAATCCGAACAACAGCAC     842

147   u...ValLeuPheGluGlyAspGluValTyrSerThrIleArgLysGlnG  163
        |||    :::|||||||||||||||||    ||:|||||
843   ATCTGTCTCCTACAATGGCCAATTATTTCAGCGACGGTC......GCCG  886

163   luTyrAsnGlyLysIleProArgPheArgArgIleArgGlyGluSerGlu  179
        :|||||||    |||  ||||||   |    ||:||||||||
887   ACTTTTCCGGGGCGACCCTCCATATACAGGAGCCCCAGCGCCACCGAA  936

180   LeuTyrThrSerAspThrValMetGlnAsn...ProGlnPheIleLysAl  195
        |||||    ||  :||::||||||||  |||||||
937   CTC......TCAGATCTCAAACAACTGAACGCACCGAATTCGTAAACTC  980

195   aThrIleValHisGlnAspIleTyrAspAspLysIleTyrTyrPheP     212
        ::||    |||||||   ||||:||||::||||
981   GGTG......GCCTATGGCGACTACATATTCTTCTTCT  1012

212   heArgGluAspAsnProAspLysAsnProGluAlaProLeuAsnValSer  228
        ::|||||    :::|||||    |||:||    |||
1013  ACCGTGAAACCCGTGCAGTACAATGCAACTGCGGAAAAGTCATCTACTCG  1062

229   ArgValAlaGlnLeuCysLysArgGlyAspTyrAspGlnGlyGluSerLeuSe  245
        |||||||||:|||||:||    :|||:|||||||||
1063  CGGGTCGCCCAGGGTGTGCAAGGACGACAAAAGGGGGCCCTCACCAG...TC  1109

245   rValSerLysTrpAsnThrPheLeuLysAlaMetLeuValCysSerAspA  262
        :|||||    |||||||||||||||||||
1110  ACGCGACCGCTGGACGTCGTTCCTCAAAGCACGTCTCAATTGTTCAATTC  1159

262   laAlaThrAsnLys..AsnPheAsnArgLeuGlnAspValPheLeuLeu  277
        :::    ||||||||||||||
1160  CCGGCGAGTACCCCTTTACTTTGATGAAATCCAATCAACAGTGATATA  1209

278   ProAspProSerGlyGlnTrpArgAspThrArg.....ValTyrGlyVa  292
        :::|||||||||    ||    :|||||||||||
1210  GTCGAGGGTCGGTACAATTCCGACGACAGCAAAAAGATCATTTATGGAAT  1259

292   lPheSerAsnProTrpAsnTyr.......SerAlaValCysValTyrS  306
        :::|||:|||    ||||||||||||:::||::
1260  CCTCACAACTCCAGTAATGCCATCGGCGTCGGCCATTTGCGCGTATC  1309

306   erLeuGlyAspIleAspLysValPheAlaArgThrSerSer.........  318
        :::::|||||   ::|||||||||   |||
1310  AAATGGCCGACATCTTGCGCGTTGAAGGGAGCTTCAAGCACCACCAAGAG  1359

319   .......... LeuLysGlyTyrHisSerSerLeuProAsnPr  329
        ||||||||   ||||||
1360  ACGATCAACTCGAACTGGCTCCCCGTGCCCCAGAACCTAGTCCCTGAACC  1409

329   oArgProGlyLysCysLeuProAspGlnProIleProThrGluThrP  346
        ||:||||||||:||||||    |:::||:
1410  CAGGGCCCGGGGCAGTCGCAGCAGCAGGACATCCTGCCCCACAAGAACG  1459

346   heGlnValAlaAspArgHisProGluValAlaGlnArgValAlaGluProMet  362
        :::    ||:::|||||||:||||||||
1460  TCAACTTTATTAAGACCACTCTTTGATGAGGACGTTCCGGCTCTTTTC  1509

363   GlyProLeuLysThrProLeuPheHisSerLysTyrHisTyrGlnLeuLysVa  379
        :::|||||||   ||||||||||||   |||
1510  GGA......AAACCAGTTCTGGTCCCAGTGAGTCTGCAGTATCGGTTTAC  1553

379   lAlaVal........HisArgMetGlnAlaSerHisGlyGluThrPheH  393
        :|||:||    |||   :::    ||:::::
1554  AGCCATAACAGTGGATCCACAAGTGAAAACAATCAATAATCAGTATCTCG  1603

393   lsValLeuTyrLeuThrThrAspArgGlyThrIleHisLysLysValValGlu  409
        |||||||||   |||||    |||   ||
1604  ATGTTTTGTATATCGGAACAGATGATGGGAAGGTACTAAAAGCTGTTAAT  1653
```

```
410   ...ProGlyGluGlnGluHisSerPheAlaPheAsnIleMetGluIleGl  425
        :|||    :::::|||   ||::|||
1654  ATACCAAAGCGACCACGCTAAAGCGTTGTTATATCGAAAATACCGTACATC  1703

425   nProPheArgArgAlaAlaAlaIleGlnThrMetSerLeuAspAlaGluA  442
        ::|||||    :||||
1704  CGTACATCCGCACGGAGCTCCCGTAAAACAGCTGAAGATCGCTCCCGGTT  1753

442   rgArgLysLeuTyrValSerSerGlnTrpGluValSerGlnValProLeu  458
        ||||||:|||    :|||:|||||||||
1754  ATGGCAAAGTTGTGGTGGTCGGGAAAGCGAAATCAGACTTGCTAATCTC  1803

459   AspLeuCysGluValTyrGlyGlyValGlyCysHisGlyCysLeuMetSerAr  475
        ::|    |||    ||||:::|||   |||:::|||:
1804  AACCAATGT..GCAAGCAAAACGCGGTGCAAGGACTGTGTGGAACTGCA  1850

475   gAspProTyrCysGlyTrpAsp......GlnGlyArgCysIleSerIle.  489
        ::||||||   :::|||   :::||||||
1851  AGACCCACATTGCGCCTGGGACGCCCAAACAAAACCTGTGTGTCAGCATTG  1900

490   ....TyrSerSerGluArgSerValLeuGlnSerIleAsnProAlaGlu  504
        :::|||    |||    ::|||||
1901  ACACCGTCACTTCGTATCGCTTCCTGATCCAGGACGTAGTTCGCGGCGAC  1950

505   ProHisLysGluCysProAsnProLysProAspLysAsnProLeuGlyLy  521
        :||    ::|||||   |||||   ||||||
1951  GACAACAA...TGTTGGTCGCCGCAAACAGACAAAAAGACTGTGATTAA  1997
```

```
521   s  521

1998  G  1998
```

seq_name: /cyn2_6/ptodata/2/lna/5B_COMB.seq:US-08-835-268-63

```
;   INFORMATION FOR SEQ ID NO: 63:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 2504 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: double
;     TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
;   FEATURE:
;     NAME/KEY: CDS
;     LOCATION: 355..2493
US-08-835-268-63

alignment_scores:
      Quality:   408.00           Length:   517
      Ratio:     1.365            Gaps:      23
Percent Similarity:  57.834       Percent Identity:  26.692

alignment_block:
US-09-041-236-2 x US-08-835-268-63

Align seg 1/1  to: US-08-835-268-63   from: 1  to: 2504
```

```
  43 HisThrValLeuPheHisGluProGlySerSerValArgValValGlyGl  59
     |||  ::::::::::::::::   ::::||||  |  ||||||||
 496 CACTTTATTGTCTGAATCAA...GACGAAACTTCGATATTAGTAGGGCCG  542

  59 yArgGlyLysValTyr........LeuPheAspPheProGluGlyLysA   73
     ||||  ::::|||   |||:::  |:::  ::::::||||||::::|||
 543 TAGAAATAGGGTTTACAATTTAGTATATTCGACCTCAGTGAGCGTAAAG  592

  73 snAlaSerValArgThrValAsnIleGlySerThrLysGlySerCysLeu  89
     :::  ::::|||   ::  ::::||::::::
 593 GGGGGCCGAATCGACTGGCCATCGTCCGATGCACATGGCCAGTTGTGTATA  642

  90 AspLys..........ArgAspCysGluAsnTyrIleThrLeuLeuGl  102
     |||             ||||||::||||||||||||::  :::|||
 643 TTGAAAGGGAAAAACGGACGACGACTGCCAAAATTACATTAGAATACTGTA  692

 102 uArgArgSerGluGly...LeuLeuAlaCysGlyThrAsnAlaArgHisP  118
     ||||||:::  |||   |||:::  ||||||:::::::||:::::
 693 CTCTTCAGAAACCGGGGAAATTAGTTATTTGCGGACCAATTCGTACAAAC  742

 118 roSerCys......TrpAsnLeuValAsnGlyThrValProLeuValGly 132
     ||  :::::::       :::||  ::::::||
 743 CCCTCTGTCGACTACGGCATTTAAGGAGGGAAAGTACCTGGTTGAGAAA   792

 133 GluMetArg......GlyTyrAlaProPheSerProAspGluAsnSerIle 147
     |||:::||||       |||  |||:::::::||||::::::||
 793 GAAGTAGAAGGGATAGGCTGTGTCCATACAATCCGGACACACAGCAC   842

 147 u...ValLeuPheGluGluGlyAspGluValTyrSerThrIleArgLysGlnG 163
     ||  :::::::  |||:::::::||||||||||||||::::::|||
 843 ATCTGTCTCCTACAATGGCCAATTATTTTCAGCGACGGTC......GCCG  886

 163 luTyrAsnGlyLysIleProArgPheArgArgIleArgGlyGluSerGlu 179
     |||  :::::::|||   |||  :::::::||||||::::||
 887 ACTTTTCCGGGGCGACCCCTCTCATATACAGGGAGCCCCAGCCACCGAA  936

 180 LeuTyrThrSerAspThrValMetGlnAsn...ProGlnPheIleLysAl 195
     |||   |||||||  |||::::::::  |||  |||::::::||||
 937 CTC.....TCAGATCTCAAACAACTGAAGCACCGACCACCACTTCGTAAACTC  980

 195 aThrIleValHisGlnAlaTyrAspAspLysIleTyrTyrProHeP 212
     :::::  ||||||:::  ||||||||::::::||||:::::
 981 GGTG..........GCCTATGGCGACTACATATTCTTTCTTCT 1012

 212 heArgGluAspAsnProAspLysAsnProGluAlaProLeuAsnValSer 228
     :::::  |||:::::||||  :::  ||::::::||||::::::
1013 ACCGTGAAACGCCGTGCAGTGACATGACAACTGCCGAAAGTCATCTACTCG 1062

 229 ArgValAlaGlnLeuCysArgGlyArgAspGlnGlyGlyGluSerLeuSe 245
     |||||||  :::||::::::  ::  ||||||
1951 GACAACAAA...TGTTGGTCGCCCAAACAGACAAAAGACTGTGATTAA 1997
```

```
1063 CGGGTCGCCAGGGTGTGCAAGGAGGACAAAGGGGGCCCTCACCAG...TC 1109

 245 rValSerLysTrpAsnThrPheLeuLysAlaMetLeuValCysSerAspA 262
     !   :::::::::||||||||:::||||||    ||  |||||||
1110 ACGGGACCCGTGACGTCGTTCCTCAAAGCACGTCCAATTGTTCAATTC 1159

 262 laAlaThrAsnLys...AsnPheAsnArgLeuGlnAspValPheLeuLeu 277
     :::                ||||||:::::::||::::::  ::
1160 CCGGCGAGTACCCCTTTACTTTGATGAAATCCAATCAACAAGTGATATA 1209

 278 ProAspProSerGlyGlnTrpArgAspThrArg......ValTyrGlyVa 292
     |||      :::||  |||||||  :::::::||:::::
1210 GTCGAGGGTGCGTACAATTCCGACGACAGCAAAAGATCATTTATGGAAT 1259

 292 lPheSerAsnProTrpAsnTyr........SerAlaValCysValTyrS 306
     :::::::|||  |||   :::::::      ||:::::  |||::::::
1260 CCTCACAACTCCAGTTAATGCCATCGGCGGCTCGGCCATTGCGCGTATC 1309

 306 erLeuGlyAspIleAspLysValPheArgThrSer............ 318
     :::::||:::  :::  |||   :::::
1310 AAATGGCGACATCTTGCGCGTGTTTGAAGGGAGCTTCAAGCACCAAGAG 1359

 319 ..........LeuLysGlyTyrHisSerSerLeuProAsnPr 329
           :::|||:::||||||::::::|||:::|||
1360 ACGATCAACTGCGACTGGCTCCCCGTGCCCAGAACCTAGTCCCTGAACC 1409

 329 oArgProGlyLysCysLeuProAspGlnGlnProIleProThrGluThrP 346
     ||||::::||||:::  ::  :::::  ::||||::::::::::
1410 CAGGCCCGGGCAGTGCTACGCGACCAGGATCCCCGACACAGAACG 1459

 346 heGlnValAlaAspArgHisProGluValAlaGlnArgValGluProMet 362
     :::  ||::::::  :::::::::||||||::::::::
1460 TCAACTTATTAAGACCCACTCTTTGATGGAGACGTTCCGGCGTCTTTC 1509

 363 GlyProLeuLysThrProLeuPheHisSerLysTyrHisTyrGlnLysVa 379
     |||          ::::::|||::::::||||::::::||
1510 GGA......AAACCAGTTCTGGTCCGAGTAGTGTCAGTATCGGTTTAC 1553

 379 lAlaVal.......HisArgMetGlnAlaSerHisGlyGluThrPheH 393
     :::::::        :::  ::::::::::
1554 AGCCATAACAGTGGATCCCACAAGTGAAAACAATCAATATCAGTATCTCG 1603

 393 isValLeuTyrLeuThrThrAspArgGlyThrIleHisLysValGlu 409
     ||::::::  :::::  ||::::::  :::
1604 ATGTTTGTATATCCGAACAGATGATGGGAAGGTACTAAAAGCTGTTAAT 1653

 410 ...ProGlyGluGlnGluHisSerPheAlaPheAsnIleMetGluIleGl 425
       |||::::::||::::::||||::::::  ::::::
1654 ATACCAAAGCGACACGCTAAAGCCGTTGTTATATCGAAATACCGTACATC 1703

 425 nProPheArgArgAlaAlaAlaAlaIleGlnThrMetSerLeuAspAlaGluA 442
     ::::::::||      |||:::::::::::::::
1704 CGTACATCCGCACGAGCCTCCCGTAAAACAGCTGAAGATCGTCCCGGTT 1753

 442 rgArgLysLeuTyrValSerSerGlnTrpGluValSerGlnValProLeu 458
     |||   :::::::||||||  :::::::||||::::::||
1754 ATGGCAAAGTGTGGTGGTGGGAAAGACGAAATCAGACTTGCTAATCTC 1803

 459 AspLeuCysGluValTyrGlyGlyGlyGlyCysHisGlyCysLeuMetSerAr 475
     :::  ||||||||||||::::||||||||||||::::::
1804 AACCATTGT...GCAAGCAAAACCGGTGCAAGGACTGTGTGGAACTGCA 1850

 475 gAspProTyrCysGlyTrpAsp.....GlnGlyArgCysIleSerIle. 489
     :::::::||||||::::::||       ||  :::::::||::::::
1851 AGACCCACATTGCGCCTGGACGCCCAAACAAAACCGTGTGTCAGCATTG 1900

 490 ....TyrSerSerGluValArgSerValLeuGlnSerIleAsnProAlaGlu 504
     ||  :::::::||||:::  ::::::||||||::::::
1901 ACACCGTACTTCGTATCGTCTTCCGATCCAGGACGTAGTTCGCGCGAC 1950

 505 ProHisLysGluCysProAsnProLysProAspLysProAlaProLeuGlnLy 521
     |||:::  ||||::  :::  ::::::::::  |
```

```
       521 s 521
       1998 |
       1998 G 1998

seq_name: /cgn2_6/ptodata/2/1na/5D_COMB.seq:US-09-060-692-63

seq_documentation_block:
;   Sequence 63, Application US/09060692
;   Patent No. 5935865
; GENERAL INFORMATION:
;   APPLICANT:  Goodman, Corey S.
;   APPLICANT:  Kolodkin, Alex L.
;   APPLICANT:  Matthes, David
;   APPLICANT:  Bentley, David R.
;   APPLICANT:  O'Connor, Timothy
;   TITLE OF INVENTION:  The Semaphorin Gene Family
;   NUMBER OF SEQUENCES:  100
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  SCIENCE & TECHNOLOGY LAW GROUP
;     STREET:  268 Bush Street, Suite 3200
;     CITY:  San Francisco
;     STATE:  CA
;     COUNTRY:  USA
;     ZIP:  94104
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Floppy disk
;     COMPUTER:  IBM PC compatible
;     OPERATING SYSTEM:  PC-DOS/MS-DOS
;     SOFTWARE:  PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/09/060,692
;     FILING DATE:
;     CLASSIFICATION:  514
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  US/08/121,713
;     FILING DATE:  13-SEP-1993
;   ATTORNEY/AGENT INFORMATION:
;     NAME:  Osman, Richard A.
;     REGISTRATION NUMBER:  36,627
;     REFERENCE/DOCKET NUMBER:  B94-002-1
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  (415)343-4341
;     TELEFAX:  (415) 343-4342
;     TELEX:
; INFORMATION FOR SEQ ID NO:  63:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH:  2504 base pairs
;     TYPE:  nucleic acid
;     STRANDEDNESS:  double
;     TOPOLOGY:  linear
;   MOLECULE TYPE:  cDNA
;   FEATURE:
;     NAME/KEY:  CDS
;     LOCATION:  355..2493
US-09-060-692-63

alignment_scores:
       Quality:  408.00              Length:     517
       Ratio:    1.365               Gaps:        23
Percent Similarity:  57.834     Percent Identity:  26.692

alignment_block:
US-09-041-236-2 x US-09-060-692-63

Align seg 1/1    to: US-09-060-692-63    from:  1    to: 2504
```

```
 43 HisThrValLeuPheHisGluProGlySerSerSerValTrpValGlyGl 59
    ||      ::::::::::   |:::::::||::: ||||||||
496 CACTTATTGTCTTGAATCAA...GACGAAACTTCGATATTAGTAGGCGG 542

 59 yArgGlyLysValTyr........LeuPheAspPheProGluGluGlyLysA 73
```

```
                        ::||||::::!|||||
543 TAGAAATAGGGTTTACAATTTAAGTATATTCGACCTCAGTGAGCGTAAAG 592

 73 snAlaSerValArgThrValAsnIleGlySerThrLysGlySerCysLeu 89
    ::::::      ||| :::::::|||::::||||          :|||
593 GGGGGCGATCGACTGGCCATCGTCCGATGCACATGGCCAGTTGTGTATA 642

 90 AspLys..........ArgAspCysGluAsnTyrIleThrLeuLeuGl 102
    |||           |||||::||::|||||||||     ::|||
643 TTGAAAGGGAAACGGACGACGACGCCCAAATTACATTAGAATACTGTA 692

102 uArgArgSerGluGly...LeuLeuAlaCysGlyThrAsnAlaArgHisP 118
    ::: |||       |:| |||||||||||||||          :::
693 CTCTTCAGAACCGGGGAAATTAGTTATTTGCGGGACCAATTCGTACAAAC 742

118 roSerCys......TrpAsnLeuValAsnGlyThrValValProLeuGly 132
    |:| ||        :::: ||:||          :::::
743 CCCTCTGTCGGACGTACGCATTTAAGGAGGGAAAGTACCTGGTTGAGAAA 792

133 GluMetArg.....GlyTyrAlaProPheSerProAspGluAsnSerLe 147
    |||:::::       ||| :::::::::||:::::::::|||||
793 GAAGTAGAAGGGATAGGCTTGGTCCATACAATCCGGAACAACAAGCAC 842

147 u...ValLeuPheGluGlyAspGluValTyrSerThrIleArgLysGlnG 163
    :::  ||||    |::::::::|||::|||||          .GCCG
843 ATCTGTCTCCTACAATGGCCAATTATTTCAGCGACGGTC......GCCG 886

163 luTyrAsnGlyLysIleProArgPheArgArgIleArgGlyGluSerGlu 179
    ::::::|||   ||||::|     ||::||||||:::|||||
887 ACTTTTCCGGGGCGACCCTCCTCATATACAGGGAGCCCCAGCGACCGAA 936

180 LeuTyrThrSerAspThrValMetGlnAsn...ProGlnPheIleLysAl 195
    |||      |||||      || :::|||     |||::::|||::|||
937 CTC......TCAGATCTCAAACAACTGAACGCACCGAATTTCGTAAACTC 980

195 aThrIleValHisGlnAspGlnAlaTyrAspAspLysIleTyrTyrPheP 212
    |||          ::  |||  :::|||||||          ||:::|
981 GGTG.........GCCTATGGCGACTACATATTCTTCTTCT 1012

212 heArgGluAspAsnProAspLysAsnProGluAlaProLeuAsnValSer 228
    ::||||||      ::: :::|||||||:::||||          |||
1013 ACCGTGAAACCGCGTCGAGTACATGAACTGCGGAAAAGTCATCTACTCG 1062

229 ArgValAlaGlnLeuCysArgGlyAspGlnGlyGlyGluSerSerLeuSe 245
    ||||||||          ::::|||            :::  |
1063 CGGGTCGCCAGGGTGTGCAAGGACGACCAAAAGGGGGCCCTCACCAG...TC 1109

245 rValSerLysTrpAsnThrPheLeuLysAlaMetLeuValCysSerAspA 262
    |  :::::|||   ::||||||          |||  ||||||
1110 ACGGGACCGCCGGTCGTTCCTCAAAGCACGTCAATTGTTCAATTC 1159

262 laAlaThrAsnLys..AsnPheAsnArgLeuGlnAspValPheLeuLeu 277
    :||||          ::: |||||||          :::
1160 CCGGCGAGTACCCCTTTTACTTTGATGAAATCCAATCAACAAGTGATATA 1209

278 ProAspProSerGlyGlnTrpArgAspThrArg......ValTyrGlyVa 292
    ::::           :::::::::|||          ::||||||
1210 GTCGAGGGTCGGTACAATTCCGACGACAGCAAAAAGATCATTTATGGAAT 1259

292 lPheSerAsnProTrpAsnTyr.........SerAlaValCysValTyrS 306
    ::::::::::|| |||   ::::::|||          |||::::|:
1260 CCTCACAACTCCAGTTAATGCCATCGCGCCGGTCGGCCATTTGCGCGTATC 1309

306 erLeuGlyAspIleAspProLysValPheArgThrSer............ 318
    ::::::::::|||   :::|||::|||||  |||
1310 AAATGGCCGACATCTTGCGCGTGTTGAAGGGAGCTTCAAGCACCAAGAG 1359

319 ..........LeuLysGlyTyrHisSerSerLeuProAsnPr 329
              |||         :::||::::::::
1360 ACGATCAACTGGCTCCCCGTGCCCCAGAACCTAGTCCCTGAACC 1409

329 oArgProGlyLysCysLeuProAspProGlnProIleProThrGluThrP 346
    ||||::::::|||          :::|||
```

1410 CAGGCCCGGCAGTGCCTACGGCACGCAGGATCCTGCCCGACAAGAACG 1459

346 heGlnValAlaAspArgHisProGluValAlaGlnArgValGluProMet 362
        ::: ||| :::   :::
1460 TCAACTTTATTAAGACCCACTCTTGATGGAGGACGTTCCGCTCTTTTC 1509

363 GlyProLeuLysThrProLeuPheHisSerLysTyrHisTyrGlnLysVa 379
        ||| ||| :::::||:::   ::
1510 GGA......AAACCAGTTCTGGTCCGACTGAGTCTGCAGATCGGTTTAC 1553

379 lAlaVal........HisArgMetGlnAlaSerHisGlyGluThrPheH 393
        :||| :::::||:::::
1554 AGCCATAACAGTGGATCCACAGTGAAAACAATCAATCAGTATCTCG 1603

393 lsValLeuTyrLeuThrThrAspArgGlyThrIleHisLysValValGlu 409
        ||||||||::: ||||| ||| |||:::|||::
1604 ATGTTTTGTATATCGGAACAGATGATGGGAAGGTACTAAAAGCTGTTAAT 1653

410 ...ProGlyGluGlnGluHisSerPheAlaPheAsnIleMetGluIleGl 425
        :::   ::::: ||| :::   :: |||
1654 ATACCAAAGCGACACGCTAAAGCGTTCTTATATCGAAAATACCGTACATC 1703

425 nProPheArgArgAlaAlaAlaIleGlnThrMetSerLeuAspAlaGluA 442
        :::::::::||||| :::::::::
1704 CGTACATCCGCACGGAGCTCCGTAAAACAGCTGAAGATCGCTCCCGGTT 1753

442 rgArgLysLeuTyrValSerSerGlnTrpGluValSerGlnValProLeu 458
        ||||||:::: |||::::   :: |||
1754 ATGGCAAAGTTGTGGTGTCGGAAGACGAAATCAGACTTGCTAATCTC 1803

459 AspLeuCysGluValTyrGlyGlyCysHisGlyCysLeuMetSerAr 475
        :::   :::::|||||
1804 AACCATTGT..GCAAGCAAAACGCGGTGCAAGGACTGTGTCGGAACTGCA 1850

475 gAspProTyrCysGlyTrpAsp.....GlnGlyArgCysIleSerIle. 489
        :|||||||||:::|| |||:::::||||||
1851 AGACCCACATTGCGCCTCGGGACGCCAAACAAACCTGTGTGTCAGCATTG 1900

490 ....TyrSerSerGluArgSerValLeuGlnSerIleAsnProAlaGlu 504
        |||| :::::
1901 ACACCGCGCACTTGCTATCGCTTCCTGATCCAGGACGACGTAGTTCGCGGCGAC 1950

505 ProHisLysGluCysProAsnProLysProAspLysProAlaProLeuGlnLy 521
        ||||| ||||| :::
1951 GACAACAAA...TGTTGGTCGCGCAAACAGACAAAAGACTGTGATTAA 1997

521 s 521
1998 G 1998

seq_name: /cgn2_6/ptodata/2/ina/PCTUS9_COMB.seq:PCT-US94-10151A-63

seq_documentation_block:
; Sequence 63, Application PC/TUS9410151A
; GENERAL INFORMATION:
;   APPLICANT: The Regents of the University of California
;   TITLE OF INVENTION: The Semaphorin Gene Family
;   NUMBER OF SEQUENCES: 66
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: FLEHR HOHBACH TEST ALBRITTON & HERBERT
;   STREET: 4 Embarcadero Center, Suite 3400
;   CITY: San Francisco
;   STATE: CA
;   COUNTRY: USA
;   ZIP: 94111-4187
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: PCT/US94/10151A
;   FILING DATE: 13-SEP-1994

CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
;   NAME: Osman, Richard A.
;   REGISTRATION NUMBER: 36,627
;   REFERENCE/DOCKET NUMBER: FP-58750-PC/RAO
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (415) 781-1989
;   TELEFAX: (415) 398-3249
;   TELEX: 910 277299 FHT UR
; INFORMATION FOR SEQ ID NO: 63:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 2504 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 355..2493
PCT-US94-10151A-63

alignment_scores:
                Quality: 408.00   Length: 517
                Ratio: 1.365   Gaps: 23
Percent Similarity: 57.834   Percent Identity: 26.692

alignment_block:
US-09-041-236-2 x PCT-US94-10151A-63

Align seq 1/1 to: PCT-US94-10151A-63   from: 1 to: 2504

43 HisThrValLeuPheHisGluProGlySerSerValTrpValGlyGl 59
        ||| :::::::::::   ::::::||| ||||||||
496 CACTTTATTGTCTTGAATCAA...GACGAAACTTCGATATTAGTAGGCGG 542

59 yArgGlyLysValTyr........LeuPheAspPheProGluGlyLysA 73
        |||:::::||| |||:::||:::
543 TAGAAAATAGGGTTTACAATTTAAGTATATTCGACCTCAGTGAGCGTAAAG 592

73 snAlaSerValArgThrValAsnIleGlySerThrLysGlySerCysLeu 89
        ::::: ||| :::||:::
593 GGGGGCGAATCGACTGGCCATCGTCCGATGCACATGGCCAGTTGTGTATA 642

90 AspLys..........ArgAspCysGluAsnTyrIleThrLeuLeuGl 102
        ||| :::|||||||| ::::||
643 TTGAAAGGGAAAACGGACGACGACTGCCAAAATTACATTAGAATACTGTA 692

102 uArgArgSerGluGly...LeuLeuAlaCysGlyThrAsnAlaArgHisP 118
        :::::||||| |||||:::||||::::
693 CTCTTCAGAACCGGCGGGAAATTAGTTATTTGCGGACCAATTCGTACAAAC 742

118 roSerCys......TrpAsnLeuValAsnGlyThrValValProLeuGly 132
        || ||| :::::
743 CCCTCTGTCGGACGTACGCATTTAAGGAGGGAAAGTACCTGGTTGAGAAA 792

133 GluMetArg......GlyTyrAlaProPheSerProAspGluAsnSerLe 147
        ||:::||| ||| ::|||||:::
793 GAAGTAGAAGGGGATAGGCTTGTGTCCATACAATCCGGAACACAACAGCAC 842

147 u...ValLeuPheGluGlyAspGluValTyrSerThrIleArgLysGlnG 163
        :::::|||||| :::||::::::
843 ATCTGTCTCCTACAATGGCCAATTATTTTCAGCAGCGGTC......GCCG 886

163 luTyrAsnGlyLysIleProArgPheArgIleArgGlyGluSerGlu 179
        :::::|||||| :::|||
887 ACTTTTCCGGGGGACCCTCTCATATACAGGGAGCCCCAGCGCACCGAA 936

180 LeuTyrThrSerAspThrValMetGlnAsn...ProGlnPheIleLysAl 195
        ||| ||:::|||||| |||:::||:::||
937 CTC........TCAGATCTCAAACAACTGAACGCCACCGAATTCGTAAACTC 980

195 aThrIleValHisGlnAlaTyrAspGlnAlaTyrAspLysIleTyrTyrPheP 212

```
 981 .::: |||||   |||   |||::::|||:||
     GGTG.............GCCTATGGCGACTACATATTCTTCTTCT 1012

 212 heArgGluAspAsnProAspLysAsnProGluAlaProLeuAsnValSer 228
     ::||||||        :::   |||
1013 ACCGTGAAACCGCCGTCGAGTACATGAACTGCGAAAAGTCATCTACTCG 1062

 229 ArgValAlaGlnLeuCysArgGlyAspGlnGlyGlyGluSerSerLeuSe 245
     ||||||||||:.::||:|||   |||:::|||||   :::  ||
1063 CGGGTCGCCCAGGGTGCCAAGGACGACAAAGGGGGCCCCTCACCAG...TC 1109

 245 rValSerLysTrpAsnThrPheLeuLysAlaMetLeuValCysSerAspA 262
     ||:::::||:::::|||   |||:||||
1110 ACGCCACCGCTGGAGCGTCGTTCCTCAAGCACGTCTCAATTGTTCAATTC 1159

 262 laAlaThrAsnLys...AsnPheAsnArgLeuGlnAspValPheLeuLeu 277
     :::||||        :::|||:::||||   :::
1160 CCGGCGAGTACCCCTTTTACTTTGATGAAATCCAATCAACAAGTGATATA 1209

 278 ProAspProSerGlyGlnTrpArgAspThrArg......ValTyrGlyVa 292
     :::  ||     ::::|||||||:::
1210 GTCCAGGGGTCGGTACAATTCCGACGACAGCAAAAAGATCATTTATGGAAT 1259

 292 lPheSerAsnProTrpAsnTyr.........SerAlaValCysValTyrs 306
     :::::||  |||        |||   |||
1260 CCTCACAACTCCAGTAATCCATCGGGCGCTCGGCCATTGCGCGTATC 1309

 306 erLeuGlyAspIleAspLysValPheArgThrSerSer.......... 318
     :::::||||:||||   ::||||||   ::
1310 AAATGGCCACATCTTGCCGCGTGTTTGAAGGAGCTTCAAGCACCAAGAG 1359

 319 .............LeuLysGlyTyrHisSerSerLeuProAsnPr 329
     |||        :::|||::||||
1360 ACGATCAACTCAGACTGGCTCCCCGTGCCCCAGAACCTAGTCCCTGAACC 1409

 329 oArgProGlyLysCysLeuProAspGlnGlnProIleProThrGluThrP 346
     ::||||:||::|||        |||   |||
1410 CAGGCCCGGGCAGTCGTACGGCAGCAGGATCCTGCCCGACAAGAACG 1459

 346 heGlnValAlaAspArgHisProGluValAlaGlnArgValGluProMet 362
     ::::|||   ||   :::   :::
1460 TCAACTTATTAAGACCCAACTCTTTGATGGAGGACGTTCCGGCTCTTTC 1509

 363 GlyProLeuLysThrProLeuPheHisSerLysTyrHisTyrGlnLysVa 379
     |||        |||   :::|||:   ::||||   ::
1510 GGA......AAACCAGTTCTGGTCCGAGTGAGTCGCAGTATCGGTTAC 1553

 379 lAlaVal........HisArgMetGlnAlaSerHisGlyGluThrPheH 393
     :||||          :::::|||   :::
1554 AGCCATAACAGTGGATCCACAAGTGAAAACAATCATAATCAGTATCTCG 1603

 393 isValLeuTyrLeuThrThrAspArgGlyThrIleHisLysValValGlu 409
     ::||||   |||        ::  :::|||   |||
1604 ATGTTTTGTATATCGGACAGATGATGGAAGGTACTAAAAGCTGTTAAT 1653

 410 ...ProGlyGluGlnGluHisSerPheAlaPheAsnIleMetGluIleGl 425
     |||   ||        :::::    :::   ::
1654 ATACCAAAGCGACACGCTAAAGCGTTGTTATATCGAAAATACCGTACATC 1703

 425 nProPheArgArgAlaAlaAlaIleGlnThrMetSerLeuAspAlaGluA 442
     :::::||||:||        :::::||||
1704 CGTACATCCGCCACCGGAGCTCCCGTAAAACAGCTGAAGATGCTCCCGGTT 1753

 442 rgArgLysLeuTyrValSerSerGlnTrpGluValSerGlnValProLeu 458
     |||   |||   |||        ||  :::
1754 ATGGCAAAGTTGTGGTGGTCGGGAAAGACCGAAATCAGACTTGCTAATCTC 1803

 459 AspLeuCysGluValTyrGlyGlyGlyCysHisGlyCysLeuMetSerAr 475
     ::  ||::::||   |||        :::
1804 AACCATTGT...GCAAGCAAAACCGGTGCAAGGACTGTGTGGAACTGCA 1850

 475 gAspProTyrCysGlyTrpAsp......GlnGlyArgCysIleSerIle. 489
     :||||||:::||||:||||   :::   ||||||
```

```
1851 AGACCCACCATTGCGCCTGGGACGCCAAACAAACCTGTGTCAGCATTG 1900

 490 .....TyrSerSerGluArgSerValLeuGlnSerIleAsnProAlaGlu 504
      :::||||        ::  :::|||::|||:::
1901 ACACCGTCACTCGTATCGCTGCTTCCTGATCCAGGACGTAGTTCGGCGAC 1950

 505 ProHisLysGluCysProAsnProLysProAspLysProLysAlaProLeuGlnLy 521
     ::  |||||||   ||||||        :::
1951 GACAACAAA...TGTTGGTCGCCGAAACAGACAAAAAGACTGTGATTAA 1997

 521 s 521
     |
1998 G 1998
```

seq_name: /cgn2_6/ptodata/2/lna/5A_COMB.seq:US-08-121-713D-61

alignment_scores:
    Quality: 406.00   Length: 670
    Ratio: 1.131   Gaps: 33
Percent Similarity: 53.582   Percent Identity: 24.478

alignment_block:
US-09-041-236-2 x US-08-121-713D-61
Align seg 1/1 to: US-08-121-713D-61   from: 1 to: 2670

```
277 LeuProAspProSerGlyGlnTrpGlnThrArgValTyrGlyValPh 293
    ||||||:||        :::||:|:::  :::||  |||||
1267 CTGCCCTCCGATAAG............AGTCGATTCTTCGCCACATT 1301

293 eSerAsnProTrpAsn........TyrSerAlaValCysValTyrSerL 307
    |:  ||||||||||||
1302 CACGACGAGCACTAATGGCCTGATTGGATCTGCCGTATGCAGTTCCACA 1351

307 euGlyAspIleAspLysValPhe................ArgThrSer 317
    ::::::|||||||:: ::||||       :|||
1352 TTAACGAGATTCAGGCTGCCTTCAATGCAAATTCAAGGAGCAATCTTCA 1401

318 Ser............LeuLysGlyTyrHisSerSerLeuProAsnProAr 330
    |||             :||||       :||||:|||
1402 TCGAATTCCGCATGGCTGCCGGTGCTTAACTCCCGGGTGCCGGAACCACG 1451

330 gProGlyLysCysLeuProAspGlnProIleProThrGluThrPheG 347
    |||||||| ||   ::|||  ::::||| :::::
1452 GCCGGGTACATGTCAACGATACATCAAACCTGCCCGATACCGTACTGA 1501

347 lnValAlaAspArgHisProGluValAlaGlnArgVal......GluPro 361
    |||||  ||| ::: ::   |||
1502 ATTTCATCAGATCCCATCCACTTATGGACAAGCCGTAAATCACGAGCAC 1551

362 MetGlyProLeu.......LysThrProLeuPheHisSerLysTyrHisTy 376
    |||                ::::|||
1552 AACAATCCAGTCTATTATAAAAGGGATTGGTCTTCACCAAGCTCGTCGT 1601

376 rGlnLysValAlaValHisArgMetGlnAlaSerHisGlyGluThrPheH 393
    ::|||::::||||||    :::     ::::
1602 TGACAAAATTCCGATTGACATCCTC............AACCAGGAATACA 1639

393 isValLeuTyrLeuThrThrAspArgGlyThrIleHisLysValValGlu 409
    |||  ||| ::| :||:|||::||||:
1640 TTGTGTACTATGTGGGCACCAATCTGGGTCGCATTTACAAAAATCGTGCAG 1689

410 ProGlyGluGlnHisSerPheAla......PheAsnIleMetGluI 424
    |||  :|||:::||:::|| |||::|:
1690 TACTACCGTAACGGAGAGTCGCGTGCCAAGCTTCTGGATATCTTCGAGGT 1739

424 eGlnProPheArgArgAlaAlaAlaAlaIleGlnThrMetSerLeuAspAlaG 441
    |||||||   ::  :||:||:::||||||:
1740 GGCTCCA............AACGAGGCCATCCAAGTGATGGAAATCAGCCAGA 1780

441 luArgArgLysLeuTyrValSerSerGlnTrpGluValSerGlnValPro 457
    ||  :::|:|||||||::  :::::::|:|||||
1781 CACGTAAGAGCCTCTACATTGGCACCGATCATCGGCATCAAGCAAATCGAC 1830

458 LeuAspLeuCysGluValTyrGlyGlyLysHisGlyCysLeuMetSe 474
    |||||:|| |:||::|:||:|
1831 CTTGGCCATGTGCAATGCGGTTACGACAACTGCTTCCGCTGCGTC..... 1875

474 rArgAspProTyrCysGlyTrpAspGlnGlyArgCysIleSerIleTyrS 491
    ||||||||||||||||||| :::  ::
1876 .CGTGATCCCTACTGCGGCTGGGATAAG...GAGGCCAATACGTGCCGAC 1921

491 erSerGluArgSerValLeuGlnSerIleAsnProAlaGluProHisLys 507
    ||   :::||:|||::                ProM
1922 CGTACGAGCTGGATTTACTGCAGGATGTG...GCCAATGAAACGAGTGAC 1968

508 GluCysProAsnProLysProAspLysProAlaProLeu...GlnLysValSe 523
    |||    ||||||:|||||:
1969 ATTTGC..........GATTCGAGTGTGCTGAAAAAGAAGATTGT 2003

523 rLeuAlaProAsnSerArgTyrTyrLeuSerCys.............ProM 536
    ::::|:           |||
2004 GGTGACCTATGGCCAGAGTGTACATCTGGGCTGTTTGCTCAAATACCCG 2053

536 etGluSerArgHisAlaThrTyrSerTrpArgHisLysGluAsnValGlu 552
    :::||  |||   |||::::|
2054 AAGTGCTGAAGAATGAGCAAGTGACCTGTATCATCACTCCAAGGACAAG 2103

553 GlnSerCysGluProGlyHisGlnSerProAsnCysIle.............. 565
```

```
28 GlyHisValGlyGlnAspArgVal...AspPheGlyGlnThrGluProHi 43
   |||  ||| |||::::||| :: ::
418 GGAAAACACGGCGCAGATCATGTCGCGGAGTTCAACTCGCGCAAGCTGTA 467

43 sThrValLeuPheHis...GluProGlySerSerValTrpValGlyG 59
   ||||| :::::::|||||||:
468 CTATCGTACAATTCCATATGAACGAAGATCGAGATACGCTCTATGTGGGAG 517

59 lyArgGlyLysValTyrLeuPheAspPhePro......GluGlyLysAsn 73
   :: |||||||:| :::||||
518 CCATGGATCGCTATTCCGTGTGAACCTGCAGAATATCTCCTCATCCAAT 567

74 AlaSerValArgThrValAsnIleGlySerThrLysGly.........Se 87
   :::::||:|||||:::|          ||||:
568 TGTAATGCGGATGCGATCAACTTGGAGCCAACACGGGATGATGTGGTTAG 617

87 rCysLeuAspLysArg............AspCysGluAsnTyrIleT 99
   |||:::::||||    ||||||:: :||:::::
618 CTGCGGTCTCCAAAGGCAAAGTCAGATCTTCGACTGCAGAAACCATGTGC 667

99 hrLeuLeuGluArgSerGluGly.....LeuLeuAlaCysGlyThr 113
   ||||||:::|:|:||   ::|||
668 GTGTCATCCAGTCAATGGACCAGGGGATAGGCTCTATGTATGCGGCACC 717

114 AsnAlaArgHisProSerCysTrpProAsnLeu...ValAsnGlyThrValVa 129
    ||||||:||   :::   ::|||||
718 AACGCCCACAATCCCAAGGATATGTTATCTATGCGGAATCTAACCCACCT 767

129 lPro...........LeuGlyGluMetArgGlyTyrAla.......P 139
    :|||
768 GCCGCGCTCGGAATATGTAATGGCGTGGGCTGTGGGCATTGCCAAGTGCC 817

139 roPheSerProAspGluAsnSer................. 146
    ::|||||
818 CCTACGATCCCCTCGCACAACTCAACTGCGATTATGTGGAGAATGGCAAT 867

146 .............................................. 146

868 CCGGGTGCTCGCGGGTTGTACTCCGGCACCAATGCGAGGAGTTCACCAA 917

147 .......LeuValLeuPheGluGlyLysAspGluValTyrSerThrIleArgL 161
        ||:  |||
918 GGCGGATACGGTTATTTTCCGCCACTGATCTGTATAATACTTCGGCTAAAC 967

161 ysGlnGluTyrAsnGlyLysLysIleProArgPheArgArgIleArgGlyGlu 177
    :: |||||||:|:|  ::
968 GTTTGGAATATAAATTCAAG..........AGGACTCTGAAATACGGAC 1005

178 SerGluLeuTyrThrSerAspThrValMetGlnAsnProGlnPheIleLeuLy 194
    |||:::: :::|||:|
1006 TCCAAGTGG...........TTGGACAAACCAAACTTTGTCGG 1037

194 sAlaThrIleValHisGlnAsnAspGlnAlaTyrAspAspLysIleTyrTyrP 211
    ::|:  ::||:::|:|||||
1038 CTCCTTTGATATT...........GGGGAGTACGGTGTATTTCT 1069

211 hePheArgGluAsnProAspAsnProGluLeuAProLeuAsnVal 227
    ||||||:::||:|
1070 TTTTCCGTGAAACCGCCTGGAATACATCAACTGCGGCAAGGCTGTCTAT 1119

228 SerArgValAlaGlnLeuCysAspArgGlyGlyLyGlyGluSerSerLe 244
    ||||||:|:||: || |
1120 TCGCGCGCATCGCACGGGGTGTGCAAGAAGGATGTGGGTGGAGAAATCTGCT 1169

244 uSerValSerLysTrpAsnThrPheLeuLysAlaMetLeuValCysSerA 261
    |||:: |||| |:||||
1170 GGCC...CACAACTGGGCGCCACCTACCCTGAAGGACCAGACTCAACTGCCAGCA 1216

261 spAlaAlaThrAsnLys...AsnPheAsnArgLeuGlnAsnAspValPheLeu 276
    ::||| |||  :::|:
1217 TCTCCCGCGCGAATTCCGTTCATTTCAACGAGGATCCAATCGGTTCTACCAG 1266
```

```
                                                      Quality:    406.00            Length:    670
                                                        Ratio:     1.131              Gaps:      33
                                      Percent Similarity:  53.582    Percent Identity:  24.478

                                      alignment_block:
                                      US-09-041-236-2 x US-08-835-268-61

                                      Align.seq 1/1 to: US-08-835-268-61  from: 1  to: 2670
```

```
                        28 GlyHisValGlyGlnAspArgVal...AspPheGlyGlnThrGluProHi 43
                           |||        ::: :::   |||    :::::  |||  :::  ||||
                       418 GGAAAACACGGCGCAGATCATGTGCGGGAGTTCAACTGCGGCAAGCTGTA 467

                        43 sThrValLeuPheHis...GluProGlySerSerValTrpValGlyG 59
                           ||||||          :::  ::  :::::  ::       ::||||||
                       468 CTATCGTACATTCCATATGAACGAAGATCGAGATACGCTCTATGTGGGAG 517

                        59 lyArgGlyLysValTyrLeuPheAspPhePro......GluGlyLysAsn 73
                           :::||||||
                       518 CCATGGATCGGCCTATTCCGTGAACCTGCAGAATATCTCCTCATCCAAT 567

                        74 AlaSerValArgThrValAsnIleGlySerThrLysGly........Se 87
                           :::::  ::: :::  :::::  ::::  :::  :::::  ::   ||
                       568 TGTAATCGGATGGCGATCAACTTGAGCCAACACGGGATGATGTGGTTAG 617

                        87 rCysLeuAspLysArg..........AspCysGluAsnTyrIleT 99
                           ||:::   :::::|||               ||||:::::::   :::
                       618 CTGCGTCTCCAAAGGCAAAGTCAGATCTTCGACTGCGAAGAACCATGTGC 667

                        99 hrLeuLeuGluArgArgSerGluGly.....LeuLeuAlaCysGlyThr 113
                           :::||||                ::: |||    ::||||||||||||
                       668 GTGTCATCCAGTCAATGGACCAGGGGGATAGCGCTCTATGTATGCGGCACC 717

                       114 AsnAlaAsrgHisProSerCysTrpAsnLeu...ValAsnGlyThrValVa 129
                           ||||||       :::   ::  ::::: |||    :::  ||| ||
                       718 AACGCCCACACAATCCCAAGGATTATGTTATCATGCGAATCTAACCCACCT 767

                       129 1Pro..........LeuGlyGlyMetArgGlyTyrAla........P
                           :||                    ||| ||| ||||||
                       768 GCCGGCCTCGGAATATGTGATTGGCGTGCTGGGCATTGCCAAGTGCC 817

                       139 roPheSerProAspGluAsnSer.................... 146
                           :::||         ||:::|||||
                       818 CCTACGATCCCCTGCAACTCAACTGCGATTTATGTGGAGAATGGCAAT 867

                       146 ......................................... 146
                       868 CCGGGGTGGTCTGCCCGGTTTGTACTCCGGCACCAATGCGGAGTTCACCAA 917

                       147 ......LeuValLeuPheGluGlyLysAspGluValTyrSerThrIleArgL 161
                           :::::: |||  :::|||:|| |||    :::::  :::::
                       918 GGCGGATACGGTTATTTTCCGCACTGATCTGTATATACTTCGGCTAAAC 967

                       161 ysGlnGluTyrAsnGlyLysIleProArgPheArgGlyIleArgGlyGlu 177
                           |||  :::::::||       ::  :::::::
                       968 GTTTGGAATATAAATTCAAG...........AGGACTCTGAAATACGAC 1005

                       178 SerGluLeuTyrThrSerAspThrValMetGlnAsnProGlnPheIleLy 194
                           ||:::           :::::::::||||::
                      1006 TCCAAGTGG.......................TTGGACAAACCAAACTTTGTCGG 1037

                       194 sAlaThrIleValHisGlnAsnAspGlnAlaTyrAspAspLysIleTyrP 211
                           :::: ||                       GGGGAGTACGGTATTTCT 1069
                      1038 CTCCTTTGATATT......................

                       211 hePheArgGluAspAsnProAspLysAsnProGluAlaProLeuAsnVal 227
                           |||||||||||          :::
                      1070 TTTTCGTGAAACGCCGTGAATACATCAACTCGCGCAAGGCTGTCTAT 1119

                       228 SerArgValAlaGlnLeuCysArgGlyLeuGlyGlyAspGlyGluSerSerLe 244
                           |||:::::: |||:::: |||:::   ::     :::   ::
                      1120 TCGGCGCATCGCACGGGTGTCGCAAGAAGGAATGTGGGTGGAAAGAATCTGCT 1169
```

```
                      2104 GGACGCTACGAGATTCGTTACTCGGCCGACCAAATACATTGAGACCACCGA 2153
                           |||                :::::         |||

                       566 ......LeuPheIleGluAsnLeuThrAlaGlnGlnTyrAlaGlnTyrHisTyrP 580
                           |||            :::::::: |||   ::: |||:::|||
                      2154 ACGTGGCCCTGGTTGTGGTTTCCGTGAACGAAGCCCATGGTGGTCGGTACG 2203

                       580 heCysGluAlaGlnGluGlySerTyrPheArgGluAlaGlnHisTrpGln 596
                           |||:::        |||||||| ::      ||||||||:::::::
                      2204 ATTGCCATTTG...GGCGGCTCGCT.TTTGTGCAGCTACAACATTACAGT 2249

                       597 LeuLeuPro..............GluAspGlyIleMetAlaGluHi 607
                           :::::::                   :::::::|||:::
                      2250 GGATGCCCACAGATGCACTCCGCGACAAGAGTAATGACTATCAGAAA 2299

                       607 sLeuLeuGly 610
                           :||||||||||
                      2300 TCTACTCGGA 2309
```

```
244  uSerValSerLysTrpAsnThrPheLeuLysAlaMetLeuValCysSerA 261
     |:::   ::||| |||::||||||||| ||| ||||||:|
1170 GGCC...CACAACTGGGCCACCTACCTGAAGGCCAGAGACTCAACTGCAGCA 1216

261  spAlaAlaThrAsnLys..AsnPheAsnArgLeuGlnAspValPheLeu 276
     ::::::  |||||||||:::::::||:||:||||:
1217 TCTCCGGCGAATTCCGTTCTATTTCAACGAGATCCAATCGGTCTACCAG 1266

277  LeuProAspProSerGlyGlnTrpArgAspThrArgValTyrGlyValPh 293
     ||||||:::  :::||  :::||||||:||
1267 CTGCCCTCCGATAAG.........AGTCGATTCTTCGCCACATT 1301

293  eSerAsnProTrpAsn.........TyrSerAlaValCysValTyrSerL 307
     :::::::  :::::::: :  ::||:||
1302 CACGACGAGCACTAATGGCCTGATTGGATCTGCCGTATGCAGTTCCACA 1351

307  euGlyAspIleAspLysValPhe............ArgThrSer 317
     :::::::||:||||:::|||          :::|||:|||
1352 TTAACGAGATTCAGGCTGCCTTCAATGGCAAATTCAAGGAGCAATCTTCA 1401

318  Ser.........LeuLysGlyTyrHisSerSerLeuProAsnProAr 330
     |||         :::|||  :::||:||||||
1402 TCGAATTCCGCATGGCTGCCGGTGCTTAACTCCCGGGTGCCGAACCACG 1451

330  gProGlyLysCysLeuProAspGlnGlnProIleProThrGluThrPheG 347
     ||||||  :::|||  ::|||  :::||:|||
1452 GCCGGGTACATGTGTCAACGATACATCAAACCTGCCCGATACCGTACTGA 1501

347  lnValAlaAspArgHisProGluValAlaGlnArgVal......GluPro 361
     :::   ||||||||  :::||  :||
1502 ATTTCATCAGATCCCATCCACTTATGGACAAAGCCGTAAATCACGAGCAC 1551

362  MetGlyProLeu.....LysThrProLeuPheHisSerLysTyrHisTy 376
     ::::|||||::  |||      |||  :::||
1552 AACAATCCAGTCTATTATAAAAGGGATTTGGTCTTCACCAAGCTCGTCGT 1601

376  rGlnLysValAlaValHisArgMetGlnAlaSerHisGlyGluThrPheH 393
     :::|||||::  |||||  ::|||  :::
1602 TGACAAAATTCGCATTGACATCCTC..........AACCAGGAATACA 1639

393  isValLeuTyrLeuThrThrAspArgGlyThrIleHisLysValValGlu 409
     ||:|||  :::|||  :::||:|||:::
1640 TTGTGTACTATGTGGGCACCAATCTGGGTCGCATTTACAAAATCGTGCAG 1689

410  ProGlyGluGlnGluHisSerPheAla......PheAsnIleMetGluIl 424
     ||||||  :::|||  :::||||||
1690 TACTACCGGTAACGGAGAGTCCGCGTTCCAAGCTTCTGGATATCTTCGAGGT 1739

424  eGlnProPheArgArgAlaAlaAlaAlaIleGlnThrMetSerLeuAspAlaG 441
     :  |||  ||||||  :::||
1740 GGCTCCA........AACGAGGCCATCCAAGTGATGGAAATCAGCCAGA 1780

441  luArgArgLysLeuTyrValSerGlnInTrpGluValSerGlnValPro 457
     ||||||||  :::|||||:::||||  :::|||
1781 CACGTAAGAGCCCTACTGCCGCTGGGATAAG...GAGGCCAATACGTGCCGAC 1830

458  LeuAspLeuCysGluValTyrGlyGlyGlyHisProHisGluProHisLysCysLeuMetSe 474
     |||  :::|||||:::||||  :::
1831 CTGGCCATGTGCAATCGACGTTACGACAACTGCTTCCGCGTC..... 1875

474  rArgAspProTyrCysGlyTrpAspAspLysGlyGlnGlyArgCysIleSerIleTyrS 491
     ||||||||  ::||  :::||||||
1876 CGTGATCCCTACTGCGGCTGGGATAAG...GATTCGAGTGTGCTGAAAAGAGATTGT 1921

491  erSerGluArgSerValLeuGlnInSerIleAsnProAlaGluProHisLys 507
     ||:  :::||||||
1922 CGTAGCAGCTGGATTTACTGCACGGATGTG..GCCAATGAAACGAGTGAC 1968

508  GluCysProAsnProLysProAspLysAlaProLeu...GlnLysValSe 523
     |||  ::::||  :::||||||
1969 ATTTGC.........GATTCGAGTGTGCTGAAAAAGAGATTGT 2003

523  rLeuAlaProAsnSerArgTyrTyrLeuSerCys.........ProM 536


2004 ::::::  :::::: |||
     GGTGACCTATGGCCAGAGTGTACATCTGGCTGTTCGTCAAAATACCCG 2053

536  etGluSerArgHisAlaThrTyrSerTrpArgHisLysGluAsnValGlu 552
     ::|||  :::|||  :  :::|||:||
2054 AAGTGCTGAAGAATGAGCAAGTGACCTGTATCATCACTCCAAGGACACAG 2103

553  GlnSerCysGluProGlyHisGlnSerProAsnCysIle......... 565
     |||      :::
2104 GGACGCTACGAGATTCGTTACTCGCCGACCAAATACATTGAGACCACCGA 2153

566  ......LeuPheIleGluAsnLeuThrAlaGlnGlnTyrGlyHisTyrP 580
     |||  :::|||  ::::::  |||:|||||
2154 ACGTGGGCCTGGTTGTGGTTCCGTGAACGAAGCCGATGGTGGTCGGTACG 2203

580  heCysGluGluGlnGluGlySerTyrPheArgGluAlaGlnHisTrpGln 596
     ::|||  |||||||  :::||:|||||
2204 ATTGCCATTTG..GGCGGCTCGCT.TTTGTGCAGCTACAACATTACAGT 2249

597  LeuLeuPro.............GluAspGlyIleMetAlaAlaGluH 607
     ||||||             :::|||
2250 GGATGCCCACAGATGCACTCCGCGCCGAACAAGAGTAATGACTATCAGAAAA 2299

607  sLeuLeuGly 610
     ||||||||||
2300 TCTACTCGGA 2309
```

us-09-041-236-2.LEE.rni

Wed Jan 19 11:15:41 2000

THIS PAGE BLANK (USPTO)

OM of: US-09-041-236-2 to: EST:*   out_format : pfs
Date: Jan 19, 2000 12:26 AM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-1998 Compugen Ltd.

Command line parameters: -DEV-xlp
-MODEL-frame+_p2n.model
-Q-/cgn2_1/USPTO.spool/US09041236/runat_15012000_164051_494/app_query.fasta.1
-DB-EST -QFMT-fasta -SUFFIX-LEE.rst -GAPOP-12.000 -GAPEXT-4.000
-MINMATCH-0.100 -LOOPCL-0.000 -LOOPEXT-0.000 -QGAPOP-4.500
-QGAPEXT-0.050 -XGAPOP-10.000 -XGAPEXT-0.500 -FGAPOP-6.000
-FGAPEXT-7.000 -YGAPOP-10.000 -YGAPEXT-0.500 -DELOP-6.000
-DELEXT-7.000 -START-1 -MATRIX-blosum62 -TRANS-human40.cd1
-LIST-45 -DOCALIGN-200 -THR_SCORE-pct -ALIGN-15 -MODE-LOCAL
-OUTFMT-pfs -NORMext -MINLEN-0 -MAXLEN-1000000 -USER-US09041236
-NCPU-6 -ICPU-3 -NO_XLPXY -WAIT -THREADS-1

Search information block:
Query: US-09-041-236-2
Query length: 634
Database: EST:*
Database sequences: 4099388
Database length: 1713965092
Search time (sec): 409.410000

score_list:

Sequence	Strd	Orig	ZScore	EScore	Len	Documentation
gb_est24:AI206011	+	812.00	1639.42	1.3e-82	459	AI206011 qg14h09.x1 Soares_plac
gb_est20:AA827795	+	791.00	1595.95	3.6e-80	476	AA827795 od08a05.s1 NCI_CGAP_GC
gb_est23:AI095332	+	715.00	1440.48	1.6e-71	469	AI095332 oy17c05.s1 Soares_sene
gb_est21:AA968218	+	705.00	1421.35	1.9e-71	410	AA968218 uh14f09.r1 Soares_mous
gb_est29:AI607683	+	692.00	1395.10	5.5e-69	410	AI607683 va91b02.y1 Soares_mous
gb_est7:W47265	-	640.00	1288.17	5.0e-63	424	W47265 zc39a12.r1 Soares_senesc
gb_est3:H02902	-	633.50	1274.21	3.0e-63	446	H02902 yj39f01.s1 Soares_placen
gb_est18:AA731451	+	595.00	1198.98	4.6e-58	336	AA731451 nz94g05.s1 NCI_CGAP_GC
gb_est11:AA260340	+	591.00	1190.37	1.4e-57	347	AA260340 va91b02.r1 Soares_mous
gb_est3:H03806	-	587.00	1179.32	5.7e-57	347	H03806 yj39f01.r1 Soares_placen
gb_est30:AI662727	-	572.50	1146.61	3.8e-55	549	AI662727 va91b02.x1 NCI_CGAP_GC
gb_est30:AA856724	+	551.00	1109.01	4.7e-53	332	AA856724 od54c01.s1 NCI_CGAP_GC
gb_est3:R33537	+	550.00	1106.51	6.5e-53	344	R33537 yh79d03.r1 Soares_placen
gb_est3:R33439	+	501.00	1007.53	2.1e-47	309	R33439 yh79d03.s1 Soares_placen
gb_est5:AU050267	+	306.00	594.62	2.1e-24	889	AU050267 Paralichthys
gb_est29:AI598158	+	301.00	589.69	4.0e-24	587	AI598158 tn14c10.x1 NCI_CGAP_Br
gb_est31:C82295	+	298.50	579.10	1.6e-23	900	C82295 C82295 Leukocyte of Japa
gb_est31:AI677829	+	265.50	517.96	3.9e-20	544	AI677829 wc80h09.x1 NCI_CGAP_Pa
gb_est38:AL118624	+	222.50	428.24	3.9e-15	619	AL118624 DKFZp761F0110_r1 761 (
gb_est19:AA747293	+	216.50	433.51	2.0e-15	157	AA747293 nx88a10.s1 NCI_CGAP_GC
gb_est3:H15319	+	212.00	409.16	4.5e-14	512	H15319 ym28c07.r1 Soares_infant
gb_est32:AI733876	+	205.00	393.66	3.3e-13	561	AI733876 zo83d09.y5 Stratagene
gb_est37:AI981751	+	203.50	388.84	6.1e-13	643	AI981751 pat.pk0064.d4.f chicke
gb_est26:AI365650	+	201.50	388.22	6.6e-13	490	AI365650 ap20h08.x1 schiller ol
gb_est32:AI733905	+	198.00	379.67	2.0e-12	546	AI733905 zo83h09.y5 Stratagene
gb_est19:AA391479	+	197.50	376.12	3.1e-12	665	AA391479 LD10146.5prime LD Dros
gb_est4:R87413	+	190.50	365.51	1.8e-11	497	R87413 ym88e10.r1 Soares_tota
gb_est10:AA165024	+	188.00	362.33	3.7e-11	427	R87413 ym88e10.r1 Soares_adult
gb_est10:AI639765	+	185.50	356.89	3.7e-11	438	AA165024 zo83d09.r1 Stratagene
gb_est17:AA646591	+	185.00	352.97	6.2e-11	549	AI639765 ms81fil.yl Soares_mous
gb_est22:AI061707	+	181.00	342.94	2.2e-10	634	AA646591 ve6d08.r1 Soares_mous
gb_est30:AI641288	-	177.50	335.16	6.0e-10	665	AI061707 LD33788.5prime LD Dros
gb_est27:AI458053	-	170.50	323.13	2.8e-09	555	AI641288 fc13all.yl Zebrafish
gb_est6:W16752	-	168.50	325.44	2.1e-09	428	AI458053 tj66h10.x1 Soares_NSF
gb_est36:AI882962	-	168.50	320.15	4.1e-09	509	W16752 zb16d03.r1 Soares_fetal
gb_est10:AA170339	-	166.00	319.11	4.7e-09	552	AI882962 fc47g07.y1 Zebrafish W
gb_est8:AA023538	-	165.50	312.05	1.2e-08	642	AA170339 ms81fil.r1 Soares_mous
gb_est25:AI258757	+	165.00	314.44	8.6e-09	492	AA023538 mh75d07.r1 Soares_mous
gb_est26:AU004634	-	162.50	310.70	1.4e-08	608	AI258757 LP02029.5prime LP Dros
gb_est29:AI555178	+	162.50	303.62	3.4e-08	709	AU004634 AU004634 Bombyx mori F
gb_est37:AI981948	+	155.00	294.30	1.1e-07	442	AI555178 UI-R-C2p-qt-f-12-0-UI
	+	152.50	293.93	1.2e-07	305	AI981948 pat.pk0069.e5.f chicke

gb_est34:AI787249   -   151.50 288.48   2.4e-07   398   ! AI787249 uj58a10.x1 Sugano m
gb_es-15:AA459837   +   146.50 275.24   1.3e-06   503   ! AA459837 zx50g12.r1 Soares_t
gb_est23:AI141990   -   146.50 276.13   1.2e-06   469   ! AI141990 oo17g10.x1 Soares_N

seq_name: gb_est24:AI206011

seq_documentation_block:
LOCUS   AI206011   459 bp   mRNA   EST   29-NOV-1998
DEFINITION qg14h09.x1 Soares_placenta_8to9weeks_2NbHP8to9W Homo sapiens cDNA
            clone IMAGE:1759953 3' similar to TR:Q64906 Q64906 SIMILAR TO
            GENBANK ACCESSION NUMBER L26081 .; mRNA sequence.
ACCESSION   AI206011
NID   g3764683
VERSION   AI206011.1 GI:3764683
KEYWORDS   EST.
SOURCE   human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 459)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   On Jan 19, 1998 this sequence version replaced gi:2151569.
            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 2686   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 446.
FEATURES   Location/Qualifiers
     source   1..459
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1759953"
                     /clone_lib="Soares_placenta_8to9weeks_2NbHP8to9W"
                     /dev_stage="two placentae; one from 8 weeks and another
                     from 9 weeks post conception"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: placenta; Vector: pT73D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; lst
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAATGGAGCGGCCGCGATTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT73 vector
                     (Pharmacia). Library constructed by Bento Soares and
                     M.Fatima Bonaldo."

BASE COUNT   87 a   134 c   151 g   87 t
ORIGIN

alignment_scores:
     Quality: 812.00   Length: 153
     Ratio: 5.377   Gaps: 0
     Percent Similarity: 98.693   Percent Identity: 98.693

alignment_block:
US-09-041-236-2 x AI206011

Aln seg 1/1: to: AI206011   from: 1 to: 459

          7 AlaAlaAlaAlaSerAlaGlnGlyHisLeuArgSerGlyProArgIlePh 23
            GCGGCCGCCGCCTCCGCCCAGGCCCACCTAAGGAGCGGACCCCGCATCTT
          1 CGGGGCCGCCGCCTCCGCCCAGGCCCACCTAAGGAGCGGACCCCGCATCTT 50

         23 eAlaValTrpLysGlyHisValGlyGlnAspArgValAspPheGlyGlnT 40
            |||||||||||||||||||||||||||||||||||||||||||||||||||
         51 CGCCGTCTGAAAGGCCATGTAGGGCAGGACCCGGGTGGACTTTGGCCAGA 100

```
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="Y"
/clone="IMAGE:1367312"
/clone_lib="NCI_CGAP_GCB1"
/tissue_type="germinal center B cell"
/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from human tonsilar cells enriched for
germinal center B cells by flow sorting (CD20+, IgD-),
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Marti (CBER).  cDNA synthesis was
primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTGAAGTGGGACGGCGGCCGCCTCATTTTTTTTTTTTTTTTTTT-
3']. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
```

```
BASE COUNT          106 a         134 c        147 g         89 t
ORIGIN
```

```
alignment_scores:
               Quality:   791.00           Length:     159
                 Ratio:     5.006             Gaps:       0
Percent Similarity:  99.371    Percent Identity:  98.742

alignment_block 1/1
US-09-041-236-2 x AA827795
```

```
Align seg 1/1   to: AA827795   from: 1    to: 476

48 HisGluProGlySerSerValTrpValGlyGlyArgGlyLysValTy  64
   CACGAGCCAGGCAGCTCCTCTGTGTGGGGGAGGACGTGGCAAGGTCTA  51
    2

64 rLeuPheAspPheProGluGlyLysAsnAlaSerValArgThrValAsnI  81
   ||||||||||||||||||||||||||||||||||||||||||||||||||
52 CCTCTTTGACTTCCCCGAGGGCAAGAACGCATCTGTGCGCACGGTGAATA  101

81 leGlySerThrLysGlySerCysLeuAspLysArgAspCysGluAsnTyr  97
   ||||||||||||||||||||||||||||||||||||||||||||||||||
102 TCGGCTCCACAAAGGGGTCCTGTCTGGATAAGCGGGACTGCGAGAACTAC  151

98 IleThrLeuLeuGluArgSerGluGlyLeuLeuAlaCysGlyThrAs  114
   |||||||||||||||||||||||||||||||||||||:::||||||||
152 ATCACTCTCCTGGAGAGGTCAGAGGGGCTGCTGGTCTGGTCTGCCACCAA  201

114 nAlaArgHisProSerCysTrpAsnLeuValAsnGlyThrValValProL  131
    ||||||||||||||||||||||||||||||||||||||||||||||||||
202 CGCCCGGCACCCCAGCTGCTGGAACCTGGTGAATGGCACTGTGGTGCA.C  250

131 euGlyGluMetArgGlyTyrAlaProPheSerProAspGluAsnSerLeu  147
    |||||||||||||||||||||||||||||.||||||||||||||||||||
251 TTGGCGGAGATGAGAGGCTACGC.CCCTTCAGCCGGACGAGAACTCCCTG  299

148 ValLeuPheGluGlyAspProValTyrSerThrIleArgLysGluGlnTy  164
    ||||||||||||||||||||||||||||||||||||||||||||||||||
300 GTTCTGTTTGAAGGGGACGAGGTGTATTCCACCATCCGGAAGCAGGAATA  349

164 rAsnGlyLysIleProArgPheArgArgGlyGluSerGluLeuT  181
    ||||||||||||||||||||||||||||||||||||||||||||
350 CAATGGGAAGATCCCTCGGTTCCGCCGCCATCCGGGGCGAGAGTGAGCTGT  399

181 yrThrSerAspThrValMetGlnAsnProGlnPheIleLysAlaThrIle  197
    ||||||||||||||||||||||||||||||||||||||||||||||||||
400 ACACCAGTGATACTGTCATGCAGAACCCACAGTTCATCAAAGCCACCATC  449

198 ValHisGlnAsnGlnAlaTyrArgAspAsp 206
    ||||||||||||||||||||||||||||||
450 GTGCACCAAGACCAGGGTTACGATGAC 476
```

```
40 hrGluProHisThrValLeuPheHisGluProGlySerSerValTrp  56
   |||||||||||||||||||||||||||||||||||||||||||||||
101 CTGAGCCGCACACGGTGCTTTCCACGAGCCAGGCAGCTCCTCTGTGTGG  150

57 ValGlyGlyArgGlyLysValLeuPheAspPheProGluGlyLysAs  73
   |||||||||||||||||||||||||||||||||||||||||||||||
151 GTGGGAGGACGTGGCAAGGTCTACCTCTTTGACTTCCCCGAGGGCAAGAA  200

73 nAlaSerValArgThrValAsnIleGlySerThrLysGlySerCysLeuA  90
   |||||||||||||||||||||||||||||||||||||||||||||||||
201 CGCATCTGTGCGCACGGTGAATATCGGCTCCACAAAGGGGTCCTGTCTGG  250

90 spLysArgAspCysGluAsnTyrIleThrLeuLeuGluArgSerGlu  106
   |||||||||||||||||||||||||||||||||||||||||||||||
251 ATAAGCGGGACTGCGAGAACTACATCACTCTCCTGGAGAGGTCAGAGTGAG  300

107 GlyLeuLeuAlaCysGlyThrAsnAlaArgHisProSerCysTrpAsnLe  123
    ||||||||||||||||||||||||||||||||||||||||||||||||||
301 GGGCTGCTGGTCTGCCACCAACGCCCGGCACCCCAGCTGCTGGAACCT  350

123 uValAsnGlyThrValValProLeuGlyGluMetArgGlyTyrAlaProP  140
    ||||||||||||||||||||||||||||||||||||||||||||||||||
351 GGTGAATGGCACTGTGGTGCCACTTGGCGGAGATGAGAGGCTATGCCCCCT  400

140 heSerProAspGluAsnSerLeuValLeuPheGluGlyAspProValTyr  156
    ||||||||||||||||||||||||||||||||||||||||||||||||||
401 TCAGTCCGGACGAGAACTACCGTGTTCTGTTTGAAGGGGACGAGGTGTAT  450

157 SerThrIle 159
    |||||||||
451 TCCACCATC 459
```

```
seq_name: gb_est20:AA827795

seq_documentation_block:
LOCUS       AA827795        476 bp    mRNA            EST       25-MAR-1998
DEFINITION  od08a05.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1367312 3'
            similar to TR:O64906 O64906 SIMILAR TO GENBANK ACCESSION NUMBER
            L26081.; mRNA sequence.
ACCESSION   AA827795
NID         g2900158
VERSION     AA827795.1  GI:2900158
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 476)
AUTHORS     NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE       National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL     Unpublished (1997)
COMMENT     On Jan 14, 1998 this sequence version replaced gi:1797284.

            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
            Ph.D., Gerald Marti, M.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Insert Length: 1737   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 431.
FEATURES             Location/Qualifiers
     source          1..476
```

seq_name: gb_est23:AI095332

seq_documentation_block:
LOCUS   AI095332   469 bp   mRNA   EST   24-SEP-1998
DEFINITION oy17c05.s1 Soares_senescent_fibroblasts_NbHSF Homo sapiens cDNA
            clone IMAGE:1666088 3' similar to TR:Q64906 Q64906 SIMILAR TO
            GENBANK ACCESSION NUMBER L26081 .; mRNA sequence.
ACCESSION   AI095332
NID   g3434308
VERSION   AI095332.1 GI:3434308
KEYWORDS   EST.
SOURCE   human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 469)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   On Jan 19, 1998 this sequence version replaced gi:2285453.
            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1645   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 468.
FEATURES   Location/Qualifiers
     source   1..469
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1666088"
                     /clone_lib="Soares_senescent_fibroblasts_NbHSF"
                     /tissue_type="senescent fibroblast"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Vector: pT7T3D (Pharmacia) with a modified
                     polylinker v_TYPE: phagemid; Site_1: Not I; Site_2: Eco
                     RI; 1st strand cDNA was primed with a Not I - oligo(dT)
                     primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCGCCGCCATTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library went through one round of
                     normalization to a Cot = 5. Library constructed by Bento
                     Soares and M.Fatima Bonaldo."
BASE COUNT   91 a   136 c   156 g   86 t
ORIGIN

alignment_scores:
     Quality: 715.00   Length: 156
     Ratio: 4.735   Gaps: 0
Percent Similarity: 96.795   Percent Identity: 96.795

alignment_block:
US-09-041-236-2 x AI095332

Align seg 1/1   to: AI095332   from: 1   to: 469

  8 AlaAlaAlaSerAlaGlnGlyHisGlyHisLeuArgSerGlyProArgIlePheAl 24
    ||||||||||||||||||||||||||||||||||||||||||||||| |||||
  4 GCCGCGCCTCCGCCCAGGCCACCTAAGGAGCGGGACCCCCATGCTTCGC 53

 24 aValTrpLysGlyHisValGlyGlnAspArgValAspPheGlyGlnThrG 41
    ||||||||||||||||||||||||||||||||||||||||||||||||||
 54 CGTCTGGAAAGGCCATGTAGGGCAGGCACCCGGTGGACTTTGCCCAGACTG 103

 41 luProHisThrValLeuPheHisGluHisGlySerSerValTrpVal 57

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
104 AGCCGCACACGGGTCTTTTCCAGAGCCAGGCAGCTCCTCTGTGTGGGTG 153

 58 GlyGlyArgGlyLysValTyrLeuPheAspPheProGluGlyLysAsnAl 74
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
154 CGAGGACGTGGCAAGGTCTACCTCTTTGACTCCCCAGAGGCAAGAAGC 203

 74 aSerValArgThrValAsnIleGlySerThrLysGlySerCysLeuAspL 91
    ||||||||||||||||||||||||||||||||||||||||||||||||||
204 ATCTGT:CGCACGTGAATATCGGCTCCACAAAGGGTCCTGTCTGGATA 253

 91 ysArgAspCysGluAsnTyrIleThrLeuLeuGluArgArgSerGluGly 107
    ||||||||||||||||||||||||||||||||||||||||||||||||||
254 AGCCGGAC:CGGAGAACTACATCACTCTYCCTGGAGAGGCGGAGTGAGGGG 303

108 LeuLeuAlaCysGlyThrAsnAlaArgHisProSerCysTrpAsnLeuVa 124
    ||||||||||||||||||||||||||||||||||||||||||||||||||
304 CTGCTGGCCTGTGGACC.AACGCCCGGCCACCCCAGCTGCTGGAACCTGT 352

124 lAsnGlyThrValProLeuGlyGluMetArgGlyTyrAlaProPheS 141
    |||||||||||||||||||||||||||||||||||||||||||||| |||||
353 GAATGGCACTGTGGTCAT.CTTGGCGAGATGAGAGGCTACGC.CCCTTCA 400

141 erProAspGluAsnSerLeuValLeuPheGluGlyAspGluValTyrSer 157
    ||||||||||||||||||||||||||||||||||||||||||||||||||
401 GCCCGGACGAGAGACTCCCCGGTTCTGTTTGAAGGGACGAGAGGTGTATTCC 450

158 ThrIleArgLysGlnGlu 163
    ||||||||||||||
451 CC.ATCCGGAAGCAGGAA 467

seq_name: gb_est21:AA968218

seq_documentation_block:
LOCUS   AA968218   422 bp   mRNA   EST   19-MAY-1998
DEFINITION uh14f09.r1 Soares mouse hypothalamus NMHy Mus musculus cDNA clone
            IMAGE:1617929 5' similar to TR:Q64906 Q64906 SIMILAR TO GENBANK
            ACCESSION NUMBER L26081 .; mRNA sequence.
ACCESSION   AA968218
NID   g3142111
VERSION   AA968218.1 GI:3142111
KEYWORDS   EST.
SOURCE   house mouse.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 422)
AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
TITLE   The WashU-HHMI Mouse EST Project
JOURNAL   Unpublished.(1996)
COMMENT   On Jan 17, 1998 this sequence version replaced gi:2043262.
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:957229
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 405.
FEATURES   Location/Qualifiers
     source   1..422
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1617929"

KEYWORDS   EST.
SOURCE   house mouse.
ORGANISM   Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
             Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 410)
AUTHORS   Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
             Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,
             Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R.,
             Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
             Waterston,R. and Wilson,R.
TITLE   The WashU-NCI Mouse EST Project 1999
JOURNAL   Unpublished (1999)
COMMENT   On Jun 22, 1998 this sequence version replaced gi:3246649.

             Contact: Marra M/WashU-NCI Mouse EST Project 1999
             Washington University School of Medicine
             4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
             Tel: 314 286 1800
             Fax: 314 286 1810
             Email: mouseest@watson.wustl.edu
             This clone is available royalty-free through LLNL ; contact the
             IMAGE Consortium (info@image.llnl.gov) for further information.
             MGI:455747
             This read is a RESEQUENCE of a previously sequenced mouse clone
             This read has been verified (found to hit its original self in the
             correct orientation)
             Possible reversed clone: similarity on wrong strand
             Seq primer: -40RP from Gibco
             High quality sequence stop: 383.

FEATURES   Location/Qualifiers
     source   1..410
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:746763"
                     /clone_lib="Soares mouse 3NME12 5"
                     /sex="unknown"
                     /tissue_type="fetus"
                     /dev_stage="12.5dpc total fetus"
                     /lab_host="DH10B"
                     /note="Organ: whole fetus; Vector: pT7T3D-Pac (Pharmacia)
                     with a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Not I - oligo(dT) primer
                     [5' TGTTACCAATCTGAAGTGGACGGCCGCCCCTTATTTTTTTTTTTTTTTTT
                     3'], on total mouse RNA (provided by Minoru Ko, Wayne
                     State Univ.]; double-stranded cDNA was ligated to Eco RI
                     adaptors (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of the modified pT7T3 vector.
                     Library went through one round of normalization, and was
                     constructed by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT   94 a   126 c   99 g   89 t
ORIGIN

alignment_scores:
     Quality: 692.00   Length: 136
     Ratio: 5.282   Gaps: 0
Percent Similarity: 96.324   Percent Identity: 93.382

alignment_block:
US-09-041-236-2 x AI607683

Align seg 1/1 to: AI607683 from: 1 to: 410

   214 GluAspAsnProAspLysAsnProGluAlaProLeuAsnValSerArgVa 230
       |||||||||||||||||||||||||||||||||||||||||||||||||||
     3 GAAGACAACCCTGACAAGAACCCCGAGGCTCCTCCAATGTGTCCCGAGT 52

   230 lAlaGlnLeuCysArgGlyAspGlnGlyGlyGluSerLeuSerValS 247
       ||||||||||||||||||||||||||||||||||||||||||||||||
    53 AGCCCAGTTGTGCAGGGGGGACCAGGGGTGGTGGAGAGTTCGTTGTCTGTCT 102

seq_name: gb_est29:AI607683

seq_documentation_block:
LOCUS   AI607683   410 bp   mRNA   EST   21-APR-1999
DEFINITION   va91b02.y1 Soares mouse 3NME12 5 Mus musculus cDNA clone
             IMAGE:746763 5' similar to TR:Q64906 Q64906 SIMILAR TO GENBANK
             ACCESSION NUMBER L26081. ;, mRNA sequence.
ACCESSION   AI607683
NID   94616850
VERSION   AI607683.1 GI:4616850

                     /clone_lib="Soares mouse hypothalamus NMHy"
                     /tissue_type="hypothalamus"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGACGGCCGCCCAAGGTTTTTTTTTTTTTTTTTTTT
                     T 3']; double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. RNA
                     provided by Dr. Wolfgang Liedtke. Library went through
                     two rounds of normalization, and was constructed by Bento
                     Soares and M.Fatima Bonaldo."

BASE COUNT   109 a   118 c   100 g   95 t
ORIGIN

alignment_scores:
     Quality: 705.00   Length: 140
     Ratio: 5.222   Gaps: 0
Percent Similarity: 96.429   Percent Identity: 90.714

alignment_block:
US-09-041-236-2 x AA968218

Align seg 1/1 to: AA968218 from: 1 to: 422

   284 TrpArgAspThrArgValThrGlyValPheSerAsnProTrpAsnTyrSe 300
       |||||||||||||||||||||||||||||||||||||||||||||||||||
     1 TGGAGAGATACCAGGTCTATGGCGTTTTCTCCAACCCCTGGAACTACTC 50

   300 rAlaValCysValTyrSerLeuGlyAspIleAspPheValPheArgThrS 317
       ||||||||||||||||||||||||||||:|||:||||||||||||||||
    51 AGCTGTCTGCGTGTATTCGCTTGGTGACATTGACAGAGTCTTCCGTACCT 100

   317 erSerLeuLysGlyTyrHisSerSerLeuProAsnProArgProGlyLys 333
       |||||||||||||||||||||||| :|||||||||||||||||||||
   101 CATCCCTCAAAGGCTACCACCACATGGGCCTTCCCAACCCTGACCCTGGCATG 150

   334 CysLeuProAspGlnGlnProIleProThrGlnThrPheGlnValAlaAs 350
       |||||| |:|||||||||||||||||||||||||||||||||||||||
   151 TGCCTCCCAAAAAGCAGCCCATACCCCACAGAAACCTTCCAGGTAGCTGA 200

   350 pArgHisProGluValAlaAlaGlnAlaArgValGluProMetGlyProLeuLysT 367
       |||||||||||||||||||||||||||||||||||||||||||||||||
   201 TAGTCACCCAGAGGTGGCTCAGAGGGGTGGAACCTATGGGCCACTGAAGA 250

   367 hrProLeuPheHisSerLysTyrHisTyrGlnLysValAlaValHisArg 383
       |||||||||||||||||||||||| :|||||:||||||||:||||||
   251 CACCATTGTTCCATTCTAAGTACCATTACCAGAAGTGGTCGTTCACCGC 300

   384 MetGlnAlaSerHisGlyGluThrPheHisValLeuTyrLeuThrThrAs 400
       ||||||||||||||||||||||||||||:||:|||||||||||||||||
   301 ATGCAAGCCAGCAATGGAGAGACCTTCCATGTGCTTTATCTAACCACAGA 350

   400 pArgGlyThrIleHisLysValValGluProGlyGluGlnHisSerP 417
       |||||||||||||||||||||||||| ||:::|||||||:|||||
   351 CAGGGGCACCATTCACAAGGTGGTGGAATCAGGGGACCAGGACCATAGCT 400

   417 heAlaPheAsnIleMetGlu 423
       ||:::||||||||||
   401 TTGTCTTCAATATCATGGAG 420

RI; 1st strand cDNA was primed with a Not I - oligo(dT)
primer(5',
TGTTACCAATCTGAAGTGGAGCGGCCGCCGATTTTTTTTTTTTTTTTTTT 3'),
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pTT3 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M.Fatima Bonaldo."

BASE COUNT   79 a.   129 c   117 g   97 t   2 others
ORIGIN

alignment_scores:
    Quality: 640.00   Length: 142
    Ratio: 4.706   Gaps: 1
    Percent Similarity: 95.775 Percent Identity: 94.366

alignment_block:
US-09-041-236-2 x W47265/rev

Align seg 1/1 to reverse of: W47265   from: 1   to: 424

103 ArgArgSerGluGlyLeuLeuAlaCysGlyThrAsnAlaArgHisProSe 119
    ||||||||||||||||||||||||||||||||||||||||||||||
424 AGGCCGGTGAGGGGTGCT.GCCTGTGGCACCAACCCCCGCACCCCAG 376

119 rCysTrpAsnLeuValAsnGlyThrValProLeuGlyGluMetArgG 136
    ||||||||||||||||||||||||||||||||| :||||||||||||
375 CTGCTGGAACCTGGTGANATGCACTGGTGTGCCACTTGGCAGAGTGA.G 327

136 lyTyrAlaProPheSerProAspGluAsn..SerLeuValLeuPheGluG 152
    || |||||||||||||||||||||||||||| :: |||||||||||
326 GAGACGCCCCCTTCAGCCCGACGAGAACGTCCCGTGGTTCGTTTTGAAG 277

152 lyAspGluValTyrSerThrIleArgLysGlnGluTyrAsnGlyLysIle 168
    ||||||||||||||||||||||||||||||| :: ||||||||||||||
276 GGGACGAGGTGTATTCCACCATCCGGAAGCAGGAATACAATGGGAAGATC 227

169 ProArgPheArgArgIleArgGlyGluSerGluLeuTyrThrSerAspTh 185
    ||||||||||||||||||||||||||||||||||||||||||||||||||
236 CCTCGGTTCCGCCGCATCCGGGGCGAGAGTGAGCTGTACACCAGTGATAC 177

185 rValMetGlnAsnProGlnPheIleLysAlaThrIleValHisGlnAspG 202
    |||||||||||||||||||||||||||||||||||||||||||||||||
176 TGTCATGCAGAACCCACAGTTCATCAAAGCCACCATCGTGCACCAGACC 127

202 lnAlaTyrAspAspLysIleTyrTyrPheArgGluAspAsnProAsp 218
    ||||||||||||||||||||||||||||||||||||||||||||||
126 AGGCTTACGATGACAAGATCTACTACTTCTTCCGAGAGGACAATCCTGAC 77

219 LysAsnProGluAlaProLeuAsnValSerArgValAlaGlnLeuCysAr 235
    ||||||||||||||||||||||||||||||||||||||||||||||
76 AAGAATCCTGAGGCTCCTCCAATGTGTCCCGTGTGGCCAGTTGTGCAG 27

235 gGlyAspGlnGlyGlyGluSerSer 243
    ||||||||||||||||||||||||
26 GGGGGCACCAGGGTGGGGAAAGTTCA 2

seq_name: gb_est3:H02902

seq_documentation_block:
LOCUS   H02902   446 bp   mRNA   EST   20-JUN-1995
DEFINITION yj39f01.s1 Soares placenta Nb2HP Homo sapiens cDNA clone
            IMAGE:151129 3', mRNA sequence.
ACCESSION   H02902
NID   g865835
VERSION   H02902.1 GI:865835
KEYWORDS   EST.
SOURCE   human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.

---

247 erLysTrpAsnThrPheLeuLysAlaMetLeuValCysSerSerAspAlaAla 263
    ||||||||||||||||||||||||||||||||||||||||||||||
103 CCAAGTGGAACACCTTCCTGAAAGCCATGTTGGTCTGCAGCGATGCAGCC 152

264 ThrAsnLysAsnPheAsnArgLeuGlnAspValPheLeuLeuProAspPr 280
    ||||||:::|||||||||||||||||||||||||||||||||||||||
153 ACCAACAGGAACTTCAATCGGCTGCAAGATGTCTTCCTGCTCCCTGACCC 202

280 oSerGlyGlnTrpArgAspThrArgValTyrArgValPheSerAsnProT 297
    ||||||||||||||||||||||||||||||||||||||||||||||||
203 CAGTGGCCAGTGGAGAGATACCAGGGTCATGGCGTTTTCCAACCCCT 252

297 rpAsnTyrSerAlaValCysValTyrSerLeuGlyValGlyAspIleAspLysVal 313
    ||||||||||||||||||||||||||||||||||||||||||||::|||
253 GGAACTACTCAGCTGTCTGCGTGTATTCCGTTGGTGACATTGACAGAGTC 302

314 PheArgThrSerSerLeuLysGlyThrTyrHisSerSerLeuProAsnProAr 330
    |||||||||||||||||||||||||||||||| :::|||||||||||||
303 TTCCGTACCTCATGGCTCAAAGGCACCTACCACATGGGCCTTCCAACCCTCG 352

330 gProGlyLysCysLeuProAspGlnGlnProIleProThrGluThrPheG 347
    |||||||||||||||||||||||| :::|||||||||||||||||
353 ACCTGNCATGTGCCTCCCAAAAAAGCAGCCCATACCCCACAGAAAACCTTCN 402

347 lnValAla 349
    |||||||
403 CAGTAGCT 410

seq_name: gb_est7:W47265

seq_documentation_block:
LOCUS   W47265   424 bp   mRNA   EST   11-OCT-1996
DEFINITION zc39a12.r1 Soares_senescent_fibroblasts_NbHSF Homo sapiens cDNA
            clone IMAGE:324670 5', mRNA sequence.
ACCESSION   W47265
NID   g1331973
VERSION   W47265.1 GI:1331973
KEYWORDS   EST.
SOURCE   human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 424)
AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
TITLE   The WashU-Merck EST Project
JOURNAL   Unpublished (1995)
COMMENT   Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 731   Std Error: 0.00
            Seq primer: mob.REGA+ET
            High quality sequence stop: 394.
FEATURES   Location/Qualifiers
     source   1..424
                     /organism="Homo sapiens"
                     /db_xref="GDB:1256182"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:324670"
                     /clone_lib="Soares_senescent_fibroblasts_NbHSF"
                     /tissue_type="senescent fibroblast"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Vector: pTT3D (Pharmacia) with a modified
                     polylinker V_TYPE: phagemid; Site_1: Not I; Site_2: Eco

REFERENCE 1 (bases 1 to 446)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 1575
High quality sequence stops: 196
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1575 Std Error: 0.00
Seq primer: Promega -21m13
High quality sequence stop: 196.

FEATURES
source 1..446
 /organism="Homo sapiens"
 /db_xref="GDB:563192"
 /db_xref="taxon:9606"
 /clone="IMAGE:151129"
 /clone_lib="Soares placenta Nb2HP"
 /sex="Female"
 /dev_stage="placenta obtained at birth (full term)"
 /lab_host="DH10B (ampicillin resistant)"
 /note="Organ: placenta; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' AACTGGAAGAATTCGCGGCCGCCAGGAATTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo. "

BASE COUNT 93 a 115 c 136 g 93 t 9 others
ORIGIN

alignment_scores:
Quality: 633.50 Length: 150
Ratio: 4.728 Gaps: 3
Percent Similarity: 89.333 Percent Identity: 86.667

alignment_block:
US-09-041-236-2 x H02902

Align seg 1/1 to: H02902 from: 1 to: 446

```
 24 AlaValTrpLysGlyHisValGlyGlnAspArgValAspPheGlyGlnTh 40
    ||||||||||||||||||||||||||||||||||||||||||||||
  1 GCCGTCTGGGAAAGGCCATGTAGGGCAGGACCGGGTGGACTTTGGCCAGAC 50

 40 rGluProHisThrValLeuPheHisGluProGlySerSerValTrpV 57
    ||||||||||||||||||||||||||||||||||||||||||||||
 51 TGAGCCGCACACGTGCTTTTCCACGAGCCAGGCAGCTCCTCTGTGTGG 100

 57 aIGlyArgGlyArgGlyLysValTyrLeuPheAspPheProGluGlyLysAsn 73
    ||||||||||||||||||||||||||||||||||||||||||||||
101 TGGGAGGACGTGCCAAGGTCACCTCTTGACTTCCCCGAGGGCCAAGAAC 150

 74 AlaSerValArgThrValAsnIleGlySerThrLysGlySerCysLeuAs 90
    ||||||||||||||||||||||||||||||||||||||||||||||
151 GCATCTGTGCCACGGTGAATATCGGCTCCACCAAAGGGGTCCTGTCTGGA 200

 90 pLysArgAspCysGluAsnTyrIleThrLeuLeuGluArgArgSerGluG 107
    ||||||||||||||||||||||||||||||||||||||||||||||
```

```
201 TAAGCGGGACTGCCAGAACTACATCACTCCTGGAGAGGCGGANTGAGG 250
107 lyLeuLeuAlaCysGlyThrArgAsnAlaArgHisProSerCysTrpAsnLeu 123
    |||||||||||||||| ||||||||||||||||||||||::::||||
251 GGCTGCTGCCCTGGACC.AACGGCCGCACCCCAGCTGCTGGAACCTG 299
124 ValAsnGlyThrValValProLeuGlyGluMetArgGlyThrAlaProPh 140
    |||||||||||||||:.||:|||||||||||||||
300 GGTGAATGCACTNTGGTCACCTTGGCCAGATGAGAGGCTACGGCCCCTTT 349
140 eSerProAspGluAsnSerLeuValLeu.PheGluGlyAspGluValTyr 156
    |||||||||||||| |||||||||| ||| |||||||||||
350 CAGCCCGACGAGAACTTCCTGGTTCNTTTNAAGGGACGAAGTTTAT 399
157 SerThrIleArgLys...GlnGluTyrAsn.GlyLysIleProArg 170
    |||||| |||   :::||||||| ||||||||||||
400 TTCANCATCCNGAAAGCAANGAATTACAATTGGGAAGATCCTTCGT 445
```

seq_name: gb_est18:AA731451

seq_documentation_block:
LOCUS AA731451 336 bp mRNA EST 23-JAN-1998
DEFINITION nz94g05.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1303160 3' similar to TR:Q64906 Q64906 SIMILAR TO GENBANK ACCESSION NUMBER L26081.; mRNA sequence.
ACCESSION AA731451
NID g2753607
VERSION AA731451.1 GI:2753607
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 336)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1407499.

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman, Ph.D., Gerald Marti, M.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 1045 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 202.

FEATURES
source 1..336
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1303160"
 /clone_lib="NCI_CGAP_GCB1"
 /tissue_type="germinal center B cell"
 /lab_host="DH10B"
 /note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from human tonsillar cells enriched for germinal center B cells by flow sorting (CD20+, IgD-), provided by Dr. Louis M. Staudt (NCI), Dr. David Allman (NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was primed with a Not I - oligo(dT) primer [5'-TGTTACCAATCTGAGTGGCACGGCCGCCCTCATTTTTTTTTTTTTTTTT-

3'). Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT   66 a   95 c   113 g   62 t
ORIGIN

alignment_scores:
    Quality: 595.00   Length: 112
    Ratio: 5.360   Gaps: 0
Percent Similarity: 99.107   Percent Identity: 97.321

alignment_block:
US-09-041-236-2 x AA731451

Align seg 1/1 to: AA731451 from: 1 to: 336

  31 GlyGlnAspArgValAspPheGlyGlnThrGluProHisThrValLeuPh 47
     |||||||||||||||||||||||||||||||||||||||||||||
   1 GGGCAGGAGACGGGGTGGACTTTGGCCAGACTGAGCCGCACGGTGCTTTT 50

  47 eHisGluProGlySerSerValTrpValGlyGlyArgGlyLysValT 64
     |||||||||||||||||||||||||||||||||||||||||||||
  51 CCACGAGCCAGGCAGCTCCCTGTGTGGGGTGGGAGGACGTGGCAAGGTCT 100

  64 yrLeuPheaspPheProGluGlyLysAsnAlaSerValArgThrValAsn 80
     |||||||||||||||||||||||||||||||||||||||||||||
 101 ACCTCTTTGACTTCCCCGAGGGCAAGAACGCATCTGTGCGCACGGTGAAT 150

  81 IleGlySerThrLysGlySerCysLeuAspLysArgAspCysGluAsnTy 97
     |||||||||||||||||||||||||||||||||||||||||||||
 151 ATCGGCTCCACAAAGGGGTCCTGTCTGGATAAGCGGGACTGCGAGAACTA 200

  97 rIleThrLeuLeuGluArgSerGluGlyLeuLeuAlaCysGlyThrA 114
     |||||||||||||||||||||||||||||||||||||||||||||
 201 CATCACTCTCCTCGGAGAGGGCGGAGTGAGGGGCTGCTGTGGCACCA 250

 114 snAlaArgHisProSerCysTrpAsnLeuValAsnGlyThrValValPro 130
     |||||||||||||||||||||||||||||||||||||||||||||
 251 ACGCCCGGCGACCCCCAGCTGCTGGAACCTGGTGAATGGCACTGTGGTGCAA 300

 131 LeuGlyGluMetArgGlyThrArgAlaProPheSerPro 142
     |||||||||||||||||||||||
 301 CTTGGCGAGATGAGAGAGCTACGCCCCCTTCAAGCCG 336

seq_name: gb_est1l:AA260340

seq_documentation_block:
LOCUS   AA260340   347 bp   mRNA   EST   18-MAR-1997
DEFINITION va91b02.r1 Soares mouse 3NME12 5 Mus musculus cDNA clone
            IMAGE:746763 5' similar to TR:G1000717 G1000717 SIMILAR TO GENBANK
            ACCESSION NUMBER L26081.; mRNA sequence.
ACCESSION   AA260340
NID   g1896824
VERSION   AA260340.1 GI:1896824
KEYWORDS   EST.
SOURCE   house mouse.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 347)
AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
TITLE   The WashU-HHMI Mouse EST Project
JOURNAL   Unpublished (1996)
COMMENT   On Sep 12, 1996 this sequence version replaced gi:1400868.

            Contact: Marra M/Mouse EST Project

WashU-HHMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:455747
Possible reversed clone: similarity on wrong strand
Seq_primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 296.

FEATURES   Location/Qualifiers
   source   1..347
                 /organism="Mus musculus"
                 /strain="C57BL/6J"
                 /db_xref="taxon:10090"
                 /clone="IMAGE:746763"
                 /clone_lib="Soares mouse 3NME12 5"
                 /sex="unknown"
                 /tissue_type="fetus"
                 /dev_stage="12.5dpc total fetus"
                 /lab_host="DH10B"
                 /note="Organ: whole fetus; Vector: pT7T3D-Pac (Pharmacia)
                 with a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                 1st strand cDNA was primed with a Not I - oligo(dT) primer
                 [5' TGTTACCAATCTGAAGTGGAGCGGCCGCCTTATTTTTTTTTTTTTTTTTTT
                 3'], on total mouse RNA [provided by Minoru Ko, Wayne
                 State Univ.]; double-stranded cDNA was ligated to Eco RI
                 adaptors (Pharmacia), digested with Not I and cloned into
                 the Not I and Eco RI sites of the modified pT7T3 vector.
                 Library went through one round of normalization, and was
                 constructed by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT   75 a   102 c   90 g   80 t
ORIGIN

alignment_scores:
    Quality: 591.00   Length: 113
    Ratio: 5.277   Gaps: 0
Percent Similarity: 99.115   Percent Identity: 96.460

alignment_block:
US-09-041-236-2 x AA260340

Align seg 1/1 to: AA260340 from: 1 to: 347

 214 GluAspAsnProAspLysAsnProGluAlaProLeuAsnValSerArgVa 230
     ||||||||||||||||||||||||||||| |||||||||||||||
   3 GAAGACAACCCTGACAAGAACCCCGAGGCTCCTCCAATGTGTCCCGAGT 52

 230 lAlaGlnLeuCysArgGlyAspGlyAspGlnGlyLySerSerLeuSerValS 247
     ||||||||||||||||| |||||||||||||||||||||||||||
  53 AGCCCAGTTGTGCAGGGGGACCAGGGYGGTGAGAGTTCGTTGTCTGTCT 102

 247 erLysTrpAsnThrPheLeuLysAlaMetLeuValCysSerAspAlaAla 263
     ||||||||||||||||| ::||||||||::|||||||||||||||
 103 CCAAGTGGAACAACCTTCCTGAAAGCCATGTTGGTCTGCAGCGATGCAGCC 152

 264 ThrAsnLysAsnPheAsnArgLeuGlnAspValPheLeuLeuProAspPr 280
     ||||:::|||||||||||||||||||||||||||| |||||||||
 153 ACCAACAGGAACTTCAATCGGCTGCAAGATGTCTTCCTGCTCCCTGACCC 202

 280 oSerGlyGlnTrpArgAspThrArgValTyrGlyValPheSerAsnProT 297
     ||||||||||||||||||||||||||||||||||||||||||||
 203 CAGTGGCCAGTGGAGAGATACCAGGGYCTATGGCGTTTTCTCCAACCCCT 252

 297 rpAsnTyrSerAlaValCysValTyrSerLeuGlyAspIleAspLysVal 313
     |||||||||||||||||||||||||||||||||:::||
 253 GGAACTACTCAGCTGTCTGCGTGTATTCGCTTGGTGACATTGACAGAGTC 302

 314 PheArgThrSerSerLeuLysGlyGlyTyrHisSerSerLeu 326
     ||||||||||||||||||||| ::::||||

```
303 TTCCGTACCTCATCGCTCAAAGGCTACCACATGGGCCTT 341
```

seq_name: gb_est3:H03806

seq_documentation_block:
LOCUS   H03806   434 bp   mRNA   EST   20-JUN-1995
DEFINITION yj39f01.r1 Soares placenta Nb2HP Homo sapiens cDNA clone
            IMAGE:151129 5' similar to SP:A49069 COLLAPSIN - ; mRNA
            sequence.
ACCESSION   H03806
NID   g866739
VERSION   H03806.1   GI:866739
KEYWORDS   EST.
SOURCE   human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 434)
AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
TITLE   The WashU-Merck EST Project
JOURNAL   Unpublished (1995)
COMMENT   Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 1575
            High qality sequence stops: 292
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Possible reversed clone: similarity on wrong strand
            Insert Length: 1575   Std Error: 0.00
            Seq primer: M13RP1
            High quality sequence stop: 292.
FEATURES   Location/Qualifiers
     source   1..434
                     /organism="Homo sapiens"
                     /db_xref="GDB:563192"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:151129"
                     /clone_lib="Soares placenta Nb2HP"
                     /sex="female"
                     /dev_stage="placenta obtained at birth (full term)"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: placenta; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     AACTGGAAGAATTCGGCGGCCGCAGGAATTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo.   "
BASE COUNT   77 a   126 c   128 g   96 t   7 others
ORIGIN

alignment_scores:
     Quality: 587.00   Length: 149
     Ratio: 4.447   Gaps: 5
     Percent Similarity: 88.591   Percent Identity: 86.577

alignment_block:
US-09-041-236-2 x H03806/rev
Align seg 1/1 to reverse of: H03806   from: 1 to: 434
```

```
336 ProAspGlnProIleProThrGluThrPhe.GlnValAlaAspArgH 352
    ||||||                       ||   :::|||||||||||||||||
433 CCCGAT........ACCCCACAAAAAACCTTCCANGTGGCTGACCGTT 393

352 is..ProGluValAla..GlnArgValGluProMetGlyProLeuLys. 366
    ||||||||:||           ::::|||||||:||||   ||:|||||||||||||
392 NAACCCGCAGAGGTGGCCCAAGAGGGTGGAANCCATGGGCCTCTGAAAG 343

367 .ThrProLeuPheHisSerLysTyrHisTyrGlnLysValAlaValHisA 383
     ||||||||||||||||||||||||||:|||||||||||||||||||||||||||
342 ACGCCCATTGTTCCACTCTAAATACCATTACCACAGAAAGTGGCCGTTCACC 293

383 rgMetGlnAlaSerHisGlyGluThrPheHisValLeuTyrLeuThrThr 399
    ||:|||||||||||||||||||||:|||||||||||||||||||||||||||||
292 GCATGCAAGCCAGCCACGGGAGACCCTTTCATGTGCTTTTACCTAACTACA 243

400 AspArgGlyThrIleHisLysValValGluProGlyGluGlnGluHisSe 416
    ||||||||||||||||||||||||||||||||||||||||||||   ::
242 GACAGGGGCCACTATCCACAAGGTGGTGAACCGGGGGAGCAGGAGCNACA 193

416 rPheAlaPheAsnIleMetGluIleGlnProPheArgArgAlaAlaAlaI 433
    |||||||||||||||||||||||||||||||||||||||||||||||||||||
192 NTTCGCCTTCAACATCATGGAGATCCAGCCTN.TCGCCGCGGCCTGCCA 144

433 leGlnThrMetSerLeuAspAlaGluArgArgLysLeuTyrValSerSer 449
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||
143 TCCAGACCATGTCGCTGGATGCTGAGCGGAGGAAGCTGTATGTGAGCTCC 94

450 GlnTrpGluValSerGlnValProLeuAspLeuCysGluValTyrGlyVal 466
    |||||||||||||||||||||||||||||||||||||||||||||||||
93  CAGTGGGAGGTGAGCCAGGTGCCCCTGGACCTGTGTGAGGTGTATGGCGG 44

466 yGlyCysHisGlyCysLeuMetSerArgArgAspProTyrCysGly 480
    |||||||||||||||||||||||||||||||||
43  GGGCTGCGCCACGGTGCCTCATGTCCGAGACCCCTACTCNGGC 1
```

seq_name: gb_est30:AI62727

seq_documentation_block:
LOCUS   AI62727   549 bp   mRNA   EST   10-MAY-1999
DEFINITION va91b02.x1 Soares mouse 3NME12 5 Mus musculus cDNA clone
            IMAGE:746763 3' similar to TR:075326 O75326 SEMAPHORIN L. ; mRNA
            sequence.
ACCESSION   AI62727
NID   g4766310
VERSION   AI62727.1   GI:476310
KEYWORDS   EST.
SOURCE   house mouse.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 549)
AUTHORS   Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
            Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,
            Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R.,
            Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
            Waterston,R. and Wilson,R.
TITLE   The WashU-NCI Mouse EST Project 1999
JOURNAL   Unpublished (1999)
COMMENT   On May 7, 1998 this sequence version replaced gi:3121474.
            Other_ESTs: va91b02.y1
            Contact: Marra M/WashU-NCI Mouse EST Project 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:455747
            This clone was previously sequenced on the 5' end only, this new
            data is from the 3' end

```
                                        600 GluAspGlyIleMetAlaGlu 606
                                            |||||||   : :|||: :
                                         63 GAGGACGGAGCCCGGCCGAC 43

seq_name: gb_est20:AA856724

seq_documentation_block:
LOCUS       AA856724      332 bp    mRNA           EST      09-MAR-1998
DEFINITION  od54c01.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1371744 3'
            similar to TR:Q64906 Q64906 SIMILAR TO GENBANK ACCESSION NUMBER
            L26081. ., mRNA sequence.
ACCESSION   AA856724
NID         g2945026
VERSION     AA856724.1  GI:2945026
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 332)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997).
COMMENT     On Jan 19, 1998 this sequence version replaced gi:2151511.

            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
            Ph.D., Gerald Marti, M.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 323.

FEATURES             Location/Qualifiers
     source          1..332
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1371744"
                     /clone_lib="NCI_CGAP_GCB1"
                     /tissue_type="germinal center B cell"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I: Site_2: Eco RI; 1st strand cDNA
                     was prepared from human tonsillar cells enriched for
                     germinal center B cells by flow sorting (CD20+, IgD-),
                     provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
                     (NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
                     primed with a Not I - oligo(dT) primer
                     [5'-TGTTACCAATCTGAAGTGGAGCGGCCGCCTCATTTTTTTTTTTTTTTTTTT-
                     3']. Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization, and was
                     constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT       66 a       95 c      109 g       62 t
ORIGIN

alignment_scores:
        Quality:    551.00                     Length:    111
        Ratio:      5.009                       Gaps:      0
Percent Similarity: 99.099     Percent Identity: 99.099

alignment_block:
US-09-041-236-2 x AA856724
```

```
FEATURES             High quality sequence stop: 323.
     source          Location/Qualifiers
                     1..549
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /map="12p"
                     /clone="IMAGE:746763"
                     /clone_lib="Soares mouse 3NME12 5"
                     /sex="unknown"
                     /tissue_type="fetus"
                     /dev_stage="12.5dpc total fetus"
                     /lab_host="DH10B"
                     /note="Organ: whole fetus; Vector: pT7T3D-Pac (Pharmacia)
                     with a modified polylinker; Site_1: Not I: Site_2: Eco RI;
                     1st strand cDNA was primed with a Not I - oligo(dT) primer
                     [5'-TGTTACCAATCTGAAGTGGAGCGGCCGCCTATTTTTTTTTTTTTTTTT
                     3'], on total mouse RNA [provided by Minoru Ko, Wayne
                     State Univ.]; double-stranded cDNA was ligated to Eco RI
                     adaptors (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of the modified pT7T3 vector.
                     Library went through one round of normalization, and was
                     constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT      102 a      146 c      155 g      144 t       2 others
ORIGIN

alignment_scores:
        Quality:    572.50                     Length:    159
        Ratio:      4.241                       Gaps:      8
Percent Similarity: 84.906     Percent Identity: 72.956

alignment_block:
US-09-041-236-2 x AI662727/rev
Align seg 1/1  to reverse of: AI662727  from: 1  to: 549

452 GluValSerGlnValProLeuAspLeuCysGluValTyrGlyGly.GlyC 468
    |||||||::::::|  |||||::  :::   ||||||  ||||  ::|
495 GAAGTGACCAGGACC...CTGGACATGGTGAGGTTCAGCGGGGCTGCAT 449

468 ySHisGlyCysLeuMetSer.ArgAspProTyrCysGlyTrpAspGlnIleGl 484
    |||                  |||:||||   |||||||||||:||
448 GC......TGCCTCATGTCCGAGA...CCTAACTGTGGCTGGGAC...... 413

484 yArgCysIle.SerIleTyr..SerSerGluArgSerValLeuGlnSer 499
    |||        ||:||:|:    |||||||||||||||||||||||||
412 .AGGACCGTGCGNTGTCTARCTCAGCTCCCAACGATAGTGTGCAGNTCC 364

500 IleAsnProAlaGluProHisLysGluCysProAsnProLysProAspLy 516
    ||||||||||||||||||::||||||::||||||||||||||||||||
363 ATTAATCCAGCGGAGCCACACAGAGAGTGTCCAACCCTAAACCAGACGA 314

516 sAlaProLeuGlnLysValSerLeuAlaProAsnSerArgTyrTyrLeuS 533
    :||||||||||||||||||::|||||||||||||||||||||||||
313 GGCTCCATTGGAGAAGGTTTCCCTGGCTCGGAACTCTCGATACTACCTGA 264

533 erCysProMetGluSerArgHisAlaThrTyrSerTrpArgHisLysGlu 549
    :|||||||||||||||||::|||||||||||:|||||::::||||
263 CCTGCCCCATGGAGTCCCGGCCACGCCCACTTACTTATGGCCGCCATGGAGGAG 214

550 AsnValGluGlnSerCysGluProGlyHisGlnSerProAsnCysIleLe 566
    ||||||||||||||||:::||||||||||||:::||||||
213 AATGTGGAACAGAGCTGTGAGCCCGGCCACCAAAGCCCTAGCTGCATCCT 164

566 uPheIleGluAsnLeuThrAlaGlnGlnLeuTyrGlyHisTyrPheCysGluA 583
    |||||||||||::|||||||::::||||||  ||||||  |||||
163 GTTCATTGAGAACCTCACGGCCCGTCAGTATGGCCCATCACCGCTGCGAGG 114

583 laGlnGluGlyLySerTyrPheArgGluAlaGlnHisTrpGlnLeuLeuPro 599
    |||||||||||||||||||::::|||||||||||::|||||||||||
113 CCCAAGAGGGCTCCTACCCCGTGAGGCTCAACACTGGGAGCTGCTGCCA  64
```

Align seg 1/1   to: AA856724   from: 1 to: 332

```
 36 AspPheGlyGlnThrGluProHisThrValLeuPheHisGluProGlySe  52
    |||||||||||||||||||||||||||||||||||||||||||||||||||
  2 GACTTGGCCAGACTGAGCCGCACACGGTGCTTTTCCACGAGCCAGGCAG    51

 52 rSerSerValGlyArgGlyGlyLysValTyrLeuPheAspProPheP     69
    ||||||||||||||||||||||||||||||||||||||||||||||||
 52 CTCCTCTGTGTGGTGGGAGGACGTGGCAAGGTCTACCTCTTTGACTTCC   101

 69 roGluGlyLysAsnAlaSerValArgThrValAsnIleGlySerThrLys  85
    ||||||||||||||||||||||||||||||||||||||||||||||||||
102 CCGAGGGCCAAGAACCATCTGCGCCACGGTGAATATCGGCTCCACAAAG   151

 86 GlySerCysLeuAspLysArgAspCysGluAsnTyrIleThrLeuLeuGl 102
    ||||||||||||||||||||||||||||||||||||||||||||||||||
152 GGGTCCTGTCTGGATAAGCGGGACTGCGAGAACTACATCACTCTCCTGA  201

102 uArgArgSerGluGlyLeuLeuAlaCysGlyThrAsnAlaArgHisPros 119
    ||||||||||||||||||||||||||||||||||||||||||||||||||
202 GAGGCGGAGTGAGGGGCTGCTGGCCTGTGGCACCAACGCACGGCACCCCA 251

119 erCysTrpAsnLeuValAsnGlyThrValValProLeuGlyGluMetArg 135
    ||||||||||||||||||||||||||||| |||||||||||||||
252 GCTGCTGGAACCTGGTGAATGGCACTGTGGTGCA.CTTGGCGGATGAGA  300

136 GlyTyrAlaProPheSerProAspGluAsnSer 146
    ||||| |||||||||||||||||||||||||||
301 GGCTACGC.CCCTTCAGCCCGGACGAGAGACTCC  332
```

seq_name: gb_est3:R33537

seq_documentation_block:
```
LOCUS       R33537         344 bp    mRNA              EST       28-APR-1995
DEFINITION  yh79d03.r1 Soares placenta Nb2HP Homo sapiens cDNA clone
            IMAGE:135941 5' similar to SP:VA39_VACCC P21062 PROTEIN ;, mRNA
            sequence.
ACCESSION   R33537
NID         g789395
VERSION     R33537.1  GI:789395
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 344)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 1078
            High quality sequence stops: 275
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Possible reversed clone; similarity on wrong strand
            Insert Length: 1078  Std Error: 0.00
            Seq primer: M13RP1
            High quality sequence stop: 275.
FEATURES             Location/Qualifiers
     source          1..344
                     /organism="Homo sapiens"
```

```
                     /db_xref="GDB:541932"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:135941"
                     /clone_lib="Soares placenta Nb2HP"
                     /sex="Female"
                     /dev_stage="placenta obtained at birth (full term)"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: placenta; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo. "
BASE COUNT     74 a    87 c    106 g    75 t    2 others
ORIGIN
```

alignment_scores:
```
     Quality:    550.00          Length:   117
     Ratio:      4.911           Gaps:     4
Percent Similarity: 95.726    Percent Identity: 94.872
```

alignment_block:
US-09-041-236-2 x R33537/rev

Align seg 1/1 to reverse of: R33537 from: 1 to: 344

```
227 ValSerArgValAlaGlnLeuCysArgGly.AspGlnGlyGlyGluSer 243
    |||||||||||||||||||||||||::::||||||||||||||||||
344 GTGTCCCGTGTGGCCNAATTGCAAGGGNGGACCAGGGTGGGGAAAGTT  295

243 erLeuSerValSer.LysTrpAsnThrPheLeu.LysAlaMetLeuVal. 258
    |||||||||||| ||||||||||||||||||||| |||||||||||||
294 CACTGTCAGTCTCCAAAGTGGAACACTTTTCGAAAAGCCATGCTGGTAA  245

259 CysSerAspAlaAlaThrAsnLysAsnPheAsnArgLeuGlnAspValPh 275
    ||||||||||||||||||||||||||||||||||||||||||||||||||
244 TGCAGTGATGCTGCCACCAACAAGAACTTCAACAGGCTGCAAGACGTCTT  195

275 eLeuLeuProSerGlyGlnTrpProSerAlaValCysValTyrGlyV 292
    ||||||||||||||||||||||||||||||||||||||||||||||
194 CCTGCTCCCTGACCCCAGCGGCCAGTGGAGGGACCACCAGGGTCATGGTG 145

292 alPheSerAsnProTrpAsnTyrSerArgAlaValCysValTyrSerLeuGly 308
    |||||||||||||||||||||||||||||||||||||||||||||||||||||
144 TTTTCTCCAACCCTGGAACTACTCAGCCCGTCTGTGTGTATTCCCTCGT   95

309 AspIleAspLysValPheArgThrSerSerLeuLysGlyTyrHisSerSe 325
    |||||||||||||||||||||||||||||||||||||||||||||||||
 94 GACATTGACAAGGTCTTCCGTACCTCCTCCCTCACTCAAGGGCTACCACTCAAG  45

325 rLeuProAsnProArgProGlyLysCysLeuProAspGlnGln 339
    ||||||||||||||||||||||||||||||||||||||||||
 44 CCTTCCCAACCCGGCGGCCCTGGCAAGTGCCTCCCCAGACCACCAG   2
```

seq_name: gb_est3:R33439

seq_documentation_block:
```
LOCUS       R33439         309 bp    mRNA              EST       28-APR-1995
DEFINITION  yh79d03.s1 Soares placenta Nb2HP Homo sapiens cDNA clone
            IMAGE:135941 3', mRNA sequence.
ACCESSION   R33439
NID         g789297
VERSION     R33439.1  GI:789297
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 309)
```

AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
           Holman,M., Hultman,M., Kucaba,T., Le M., Lennon,G., Marra,M.,
           Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
           Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
           Wilson.R.
TITLE   The WashU-Merck EST Project
JOURNAL   Unpublished (1995)
COMMENT   Contact: Wilson RK
           Washington University School of Medicine
           4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
           Tel: 314 286 1800
           Fax: 314 286 1810
           Email: est@watson.wustl.edu
           Insert Size: 1078
           High quality sequence stops: 273
           Source: IMAGE Consortium, LLNL
           This clone is available royalty-free through LLNL ; contact the
           IMAGE Consortium (info@image.llnl.gov) for further information.
           Possible reversed clone: polyT not found
           Insert Length: 1078   Std Error: 0.00
           Seq primer: Promega -21m13
           High quality sequence stop: 273.
FEATURES   Location/Qualifiers
     source   1..309
                   /organism="Homo sapiens"
                   /db_xref="GDB:541932"
                   /db_xref="taxon:9606"
                   /clone="IMAGE:135941"
                   /clone_lib="Soares placenta Nb2HP"
                   /sex="female"
                   /dev_stage="placenta obtained at birth (full term)"
                   /lab_host="DH10B (ampicillin resistant)"
                   /note="Organ: placenta; Vector: pT7T3D (Pharmacia) with a
                   modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                   strand cDNA was primed with a Not I - oligo(dT) primer [5'
                   AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTTT 3'],
                   double-stranded cDNA was ligated to Eco RI adaptors
                   (Pharmacia), digested with Not I and cloned into the Not I
                   and Eco RI sites of the modified pT7T3 vector. Library
                   went through one round of normalization. Library
                   constructed by Bento Soares and M.Fatima Bonaldo.   "
BASE COUNT   58 a   86 c   104 g   58 t
ORIGIN

alignment_scores:
     Quality: 501.00   Length: 101
     Ratio: 5.165   Gaps: 1
Percent Similarity: 96.040   Percent Identity: 95.050

alignment_block:
US-09-041-236-2 x R33439   ..
Align seq 1/1 to: R33439   from: 1 to: 309

    11 SerAlaGlnGlyHisLeuArgSerGlyProArgIlePheAlaValTrpLy   27
        ::|||||||||||||||||||||||||||||||| ||||||||||
     7 GCCGCCCAGGGCCACCTAAGGAGCGGGATNCTANNTCTTCGCCGTCTGGAA   56

    27 sGlyHisValGlyValGlyGlnAsnAspArgValAspPheGlyGlnThrGluProHisT   44
        ||||||||||||||||||||||||||||||||||| ||||||||||
    57 AGGCCATGTAGGGCAGGACCGGGTGGACTTTGGCCAGACTGAGCCGCCACA   106

    44 hrValLeuPheHisGluProGlySerSerValTrpValGlyGlyArg   60
        |||||||||||||||||||||||||||||||||||||||||||||||
    107 CGGTGCTTTTCCAGAGCCAGGCAGCTCCTCTGTGTGGGTGGGAGGACGT   156

    61 GlyLysValTyrLeuPheAspPheProGluGlyLysLysAsnAlaSerValAr   77
        |||||||||||||||||||||||||||||||||||||||||||||||
    157 GGCAAGGTCTACCTCTTTGACTTCCCGAGGGCAAGAACGCATCTGTGCG   206

    77 gThrValAsnIleGlySerThrLysGlySerCysLeuAspLysArgAspC   94
        |||||||||||||||||||||||||||||||||||||||||||||||

           ||||||||||||||||||||||||||||||||||||||||||||||||||
    207 CACGGTGAATATCGGCTCCACAAAGGGTCCGTCTGGATAAGCGGGACT   256

    94 ysGluAsnTyrIleThrLeuLeu.GluArgArgSerGluGlyLeuLeuAl   110
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    257 GCGAGAACTACATCACTCTCCTGGAGAGGCGGGAGTGAGGGGCTGCTGGC   306

    110 a 110
        |
    307 C 307

seq_name: gb_est32:AU050267

seq_documentation_block:
LOCUS   AU050267   889 bp   mRNA   EST   08-JUN-1999
DEFINITION AU050267 Paralichthys olivaceus library (Aoki T) Paralichthys
            olivaceus cDNA clone WD6-13, mRNA sequence.
ACCESSION   AU050267
NID   g5018720
VERSION   AU050267.1 GI:5018720
KEYWORDS   EST.
SOURCE   flounder.
  ORGANISM Paralichthys olivaceus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
            Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Percomorpha;
            Pleuronectiformes; Pleuronectoidei; Bothidae; Paralichthys.
REFERENCE   1 (bases 1 to 889)
  AUTHORS   Aoki,T.
  TITLE   Japanese flounder leukocyte cDNA
  JOURNAL   Unpublished (1999)
COMMENT   On Dec 20, 1995 this sequence version replaced gi:1131417.
            Contact: Takashi Aoki
            Genetics and Biochemistry
            Tokyo University of Fisheries
            Konan, 4-5-7, Minato 108, Japan
            Tel: 03-5463-0556
            Fax: 03-5463-0690
            Email: ad95217@s4201.tokyo-u-fish.ac.jp
            clone WD6-13:similar to mouse semaphorin C(X85992).
FEATURES   Location/Qualifiers
     source   1..889
                   /organism="Paralichthys olivaceus"
                   /db_xref="taxon:8255"
                   /map="5"
                   /clone="WD6-13"
                   /clone_lib="Paralichthys olivaceus library (Aoki T)"
BASE COUNT   212 a   217 c   249 g   206 t   5 others
ORIGIN

alignment_scores:
     Quality: 306.00   Length: 285
     Ratio: 1.877   Gaps: 14
Percent Similarity: 57.193   Percent Identity: 32.632

alignment_block:
US-09-041-236-2 x AU050267
Align seq 1/1 to: AU050267   from: 1 to: 889

    268 PheAsnArgLeuGlnAspValPheLeuLeuProAspProSerGly...Gl   283
        ||||| |||||||||||||||||||||||||| || ::
      7 TTCAACATTCTCCAGGACGTGTTCACTCTGCAGCCACCAGAGGGCAGCAA   56

    283 nTrpArgAspThrArgValTyrGlyValPheSerAsnProTrpAsn....   298
        :|||||| |||||||||||||||||| |||:::
     57 TCCCCTGCACACCCGTCTCATGGCGTCTCACCTCCAGTGGTCCTCAG   106

    299 ....TyrSerAlaValCysValTyrSerLeuGlyAspIleAspLysVal   313
        ||||||||||||||||||||||||||||||||||||
    107 GTCGTCAGTCGTCGTGTCGTGCGTCTTTAAGCTGCAGGACATCAGGACCGTG   156

```
314 Phe..........ArgThr.........................SerSe 318
       |||        ||||||                        |||||
157 TTCTCTGGGAGCTATAGGACCTTCGACATGGAGATGCCACCAGTGGAGTTC 206

318 rLeuLysGlyTyrHisSerSerLeuProAsnProArgProGlyLysCysL 335
    ||||  |||  ||||||| |||||  |||                 ||:::|||
207 ACTGCTGGGAAAACACTCGTACCTG..............GGACAGTGTG 241

335 euProAspGlnGlnProIleProThrGluThrPheGlnValAlaAspArg 351
       |||                        :::::::::||||
242 GACTGGAC.........................AAAGCCACCGATTCA 264

352 HisProGluValAlaGlnArg.....................ValGluPr 361
    :::  |||  ||| :::::|||                    |||::|||
265 GAGCTAGAGGAGGTGAAGAGGAGTTTCTGACCAGTGGTGGAGTCAAACC 314

361 oMet......GlyProLeuLysThrProLeuPheHisSerLysTyrHisT 376
    |::::     ||||||:::          |||::::::::::::|
315 TGTGGCAGACGGTCCAGTG.........GTTGTTCCTCTGAGCATCAGT 355

376 yrGlnLysValAlaValHisArgMetGlnAlaSerHisGlyGluThrPhe 392
    ||:::::|||:||| |||  ||||||::||| :::: :::
356 ACAGTCGTGTGGGTGCGGATGAGGACGCAGCAAACGGCAAACAGTAC 405

393 HisValLeuTyrLeuThrThrAspArgGlyThrIleHisLysValVal.. 408
    |||:::|||:|||:::||| ::: :::||||||::::::::|||
406 ACCGTCCTTTTTCTGCNCACAGAGTCTGGATTCCTGCACAAGATTGTTGT 455

409 ....GluProGlyGluGlnGluHisSerPheAlaPheAsnIleMetGluI 424
        |||  |||                 ||| ||||  ||||
456 TCTGGATTGGGGT.....................GCCCTGGTCATTGAGGAGA 487

424 leGlnProPheArgArgAlaAlaAlaAlaIleGlnThrMetSerLeuAspAla 440
    ||||| :::     ::::::::::  |||:::::: |||::::::
488 TTCAGGTCTTCACACAGCCACAGCTGGTGAAGAGCATCGTTCTCCGACC 537

441 GluArgArgLysLeuTyrValSerSerGlnTrpGluValSerGlnValPr 457
    ||||||||  :::|||     :::|||     |||::::::|||||||
538 ACCAGGGGGGCAGTATACGNGGAACGTCTGAGGGTGTCACCTCCGTCCC 587

457 oLeuAspLeuCysCysHisGluValTyrGlyGlyLyGlyCysHisGlyCysLeuMetS 474
    |::::  ||||||||:::::||| ::::||| ||:::||||||||||::::::::::
588 TGTGGCGAACTGCTCCGTTAC...AGAAGCTGCGGCAGTGTGTTCTGG 634

474 erArgAspProTyrCysGlyTrpAspGlnGlyArg......CysIleSer 488
    ::||||||||| |||::::::     ||| |||  |||  :::
635 CCAGAGATCCTCGTGTGGGTGGGGTCGACCAGGAGGGAGTGCACAAAT 684

489 IleTyrSerSerGluArgSerValLeuGlnSerIleAsnProAlaGluPr 505
    :::  ||||||||| :::||| |||||:::::::
685 CTGAACAACGATCGNGAACACATAGNTCAGGATGTGGAA............. 723

505 oHisLysGluCysProAsnProLysProAspLysProLeuGlnLysV 522
    |||   :::|||||| |||||||| ||||||||:::::
724 .............AATGGAAATATAGAAAAAGCATGTCTACAACAAA 757

522 alSer 523
    ::|||
758 TCAGT 762
```